Galectin 11

This is a continuation-in-part of Application Serial No. 09/010,146, filed January 21, 1998 which claim the benefit of the filing date of the Provisional Application Serial Nos. 60/034,204 filed January 21, 1997 and 60/034,205 filed January 21, 1997, each of which is herein incorporated by refcrence in its entirety.

Background of the Invention

Field of the Invention

The present invention relates to a novel galectin. More specifically, isolated nucleic acid molecules are provided encoding human galectin 11. Galectin 11 polypeptides are also provided, as are vectors, host cells, recombinant methods for producing the same, and antibodies to galectin 11 polypeptides. The invention further relates to screening methods for identifying agonists and antagonists of galectin 11 activity. Also provided are diagnostic methods for detecting cell growth disorders and therapeutic methods for cell growth disorders, including autoimmune diseases, cancer, and inflammatory diseases.

Related Art

Lectins are proteins that bind to specific carbohydrate structures and can thus recognize particular glycoconjugates. Barondes et al., J. Biol. Chem. 269(33):20807-20810 (1994). Galectins are membcrs of a family of β-galactoside-binding lectins with related amino acid sequences (For review see, Barondes et al., Cell 76:597-598 (1994); Barondes et al., J. Biol. Chem. 269(33):20807-20810 (1994)). Although a large number of glycoproteins containing β-galactoside sugars are produced by the cell, only a few will bind to known galectins *in vitro*. Such apparent binding specificity suggests a highly specific functional role for the galectins.

Galectin 1 (conventionally termed *LGALS1* for lectin, galactoside-binding, soluble -1, but which is also known as: L-14-1, L-14, RL-14.5, galaptin, MGBP, GBP, BHL, CHA, HBP, HPL, HLBP 14, rIML-1) is a homodimer with a subunit molecular mass of 14,500 Daltons. Galectin 1 is expressed abundantly in smooth and skeletal muscle, and to a lesser extent in many other cell types (Couraud et al., J. Biol. Chem. 264:1310-1316 (1989). Galectin 1 is thought to specifically bind laminin, a highly polylactosaminated cellular glycoprotein, as well as the highly

polylactosaminated lysosome-associated membrane proteins (LAMPs). Galectin 1 has also been shown to bind specifically to a lactosamine-containing glycolipid found on olfactory neurons and to integrin a_7b_1 on skeletal muscle cells.

Other members of the Galectin family have also been reported. Galectin 2 was originally found in hepatoma and is a homodimer with a subunit molecular mass of 14,650 Daltons (Gitt et al., J. Biol. Chem. 267:10601-10606 (1992)). Galectin 3 (a.k.a., Mac-2, EPB, CBP-35, CBP-30, and L-29) is abundant in activated macrophages and epithelial cells and is a monomer with an apparent molecular mass between 26,320 and 30,300 Daltons (Cherayil et al., Proc. Natl. Acad. Sci. USA 87: 7324-7326 (1990)). Galectin 3 has been observed to bind specifically to laminin, immunoglobulin E and its receptor, and bacterial lipopolysaccharides. Galectin 4 has a molecular mass of 36,300 Daltons and contains two carbohydrate-binding domains within a single polypeptide chain (Oda et al., J. Biol. Chem. 268:5929-5939 (1993)). Galectins 5 and 6 are discussed in Barondes et al., Cell 76:597-598 (1994). Human Galectin 7 has a molecular mass of 15,073 Daltons and is found mainly in stratified squamous epithelium (Madsen et al., J. Biol. Chem. 270(11):5823-5829 (1995)).

Animal lectins, in general, often function in modulating cell-cell and cell-matrix interactions. Galectin 1 has been shown to either promote or inhibit cell adhesion depending upon the cell type in which it is present. Galectin 1 inhibits cell-matrix interactions in skeletal muscle presumably, by galectin 1-mediated disruption of laminin-integrin a_7b_1 interactions (Cooper et al., J. Cell Biol. 115:1437-1448 (1991)). In several non-skeletal muscle cell types, Galectin 1 promotes cell-matrix adhesion possibly by cross-linking cell surface and substrate glycoconjugates (Zhou et al., Arch. Bioch. Biophys. 300:6-17 (1993); Skrincosky et al., Cancer Res. 53:2667-2675 (1993)).

Galectin 1 also participates in regulating cell proliferation (Wells et al., Cell 64:91-97 (1991)) and some immune functions (Offner et al., J. Neuroimmunol. 28:177-184 (1990)). Galectin 1 induces the release of tumor necrosis factor from macrophages (Kajikawa et al., Life Sci. 39:1177-1181 (1986). Galectin 1 has also been demonstrated to have therapeutic activity against autoimmune diseases in animal models for experimental myasthenia gravis, and experimental autoimmune encephalomyelitis (Levi et al., Eur. J. Immunol. 13:500-507 (1983); and Offner et al., J. Neuroimmunol. 28:177-184 (1990), respectively). Additionally, galectin 1 has been shown to regulate immune response by mediating apoptosis of T cells (Perillo et al., Nature 378:736-739 (1995)).

Galectin 3 promotes the growth of cells cultured under restrictive culture conditions (Yang et al., Proc. Natl. Acad. Sci. USA 93:6737-6742 (June 1996)).

Galectin 3 expression in cells confers resistance to apoptosis which indicates that galectin 3 could be a cell death suppresser which interferes in a common pathway of apoptosis. *Id.* Galectin 3 has also been observed to function in modulating cell-adhesion, as well as in the activation of certain immune cells by cross-linking IgE and IgE receptors.

Recently, a galectin-like antigen designated HOM-HD-21 was found to be highly expressed in a Hodgkin's Disease cDNA library and another galectin, termed PCTA-1, was identified as a specific cell surface marker on human prostate cancer cell lines and patient-derived carcinomas.

Thus, galectins have been observed to be involved in the regulation of immune cell activity, as well as in such diverse processes as cell adhesion, proliferation, inflammation, autoimmunity, and metastasis of tumor cells. Accordingly, there is a need in the art for the identification of novel galectins which can serve as useful tools in the development of therapeutics and diagnostics for regulating immune response, inflammatory diseasc and cancer.

Summary of the Invention

The present invention provides isolated nucleic acid molecules comprising a polynucleotide encoding the galectin 11 polypeptide having the amino acid sequence shown in Figure 1 (SEQ ID NO:2), the amino acid sequence encoded by the cDNA clone deposited in a bacterial host as ATCC Deposit Number 209053, on May 16, 1997, and fragments, variants, derivatives, and analogs thereof.

The present invention also relates to recombinant vectors which include the isolated nucleic acid molecules of the invention, and to host cells containing the recombinant vectors, as well as to methods of making such vectors and host cells and for using them for production of galectin 11 polypeptides by recombinant techniques.

The invention further provides isolated galectin 11 polypeptide having an amino acid sequence encoded by a polynucleotide described herein and antibodies which bind these polypeptides.

The present invention also provides screening methods for identifying compounds capable of enhancing or inhibiting a cellular response, such as, for example, apoptosis, induced by galectin 11. Generally, these methods involve contacting galectin 11, the candidate compound, and a cell which expresses a galectin 11 ligand, assaying a cellular response resulting from the binding of galectin 11 with the ligand, and comparing the cellular response to a standard, the standard being assayed when contact of galectin 11 and the galectin 11 ligand is made in the absence of

the candidate compound; whereby, an increased cellular response over the standard indicates that the compound is an agonist and a decreased cellular response over the standard indicates that the compound is an antagonist.

In another aspect, a screening assay for agonists and antagonists is provided which involves determining the effect a candidate compound has on galectin 11 binding to a β-galactoside sugar. In particular, the method involves contacting a β-galactoside sugar with a galectin 11 polypeptide and a candidate compound and determining whether galectin 11 binding to the β-galactoside sugar is increased or decreased due to the presence of the candidate compound.

The invention also provides diagnostic methods useful during diagnosis of disorders associated with elevated, decreased, or otherwise aberrant expression of galectin 11.

The invention further provides for methods for treating an individual in need of an increased level of galectin 11 activity in the body comprising, administering to such an individual a composition comprising a therapeutically effective amount of an isolated galectin 11 polypeptide, fragment, variant, derivative, or analog of the invention, or an agonist thereof.

In another embodiment, the invention provides for methods for treating an individual in need of a decreased level of galectin 11 activity in the body comprising, administering to such an individual a composition comprising a therapeutically effective amount of a galectin 11 fragment, variant, derivative, analog or antibody of the invention or galectin 11 antagonist.

Brief Description of the Figures

Figure 1 shows the nucleotide sequence (SEQ ID NO:1) and deduced amino acid sequence (SEQ ID NO:2) of galectin 11. The protein has a deduced molecular mass of about 14.8 kDa.

Figure 2 shows the regions of similarity between the amino acid sequences of the galectin 11 protein (HJACE54), rat galectin 5 (SEQ ID NO:3), and human galectin 8 (SEQ ID NO:4). Identical amino acids shared between the galectins are shaded, while conservative amino acid changes are boxed.

Figure 3 shows structural and functional features of galectin 11 (SEQ ID NO:2) predicted using the default parameters of the indicated computer programs. Alpha, beta, turn and coil regions; hydrophilicity and hydrophobicity; amphipathic regions; flexible regions; antigenic index and surface probability are shown. In the Antigenic Index - Jameson-Wolf graph, amino acid residues 65-70 and 118-124 in Figure 1

(SEQ ID NO:2) correspond to the shown highly antigenic regions of the galectin 11 polypeptide.

Figure 4. Structure of human galectin 11 gene. The human galectin 11 gene is located on chromosome 11. This figure shows the structure of the region of chromosome 11 containing the galectin 11 genc and discloses the number of nucleotides corresponding to the transcribed (shaded) and untranscribed (open) portions of this region of the chromosome. The human galectin 11 gene contains 5 exons. The translation initiation site is located on the second exon. The nucleotide numbering identified in cxons designated by roman numerals correspond to that presented in Figure 1 (SEQ ID NO:1).

Figure 5A is a bar graph showing that transfection of Jurkat cells with a galectin 11 expression construct (pEF-Leg11) induces apoptosis of transfected cells. Shaded bars represent % apoptosis of Jurkat cells that have been transfected with the galectin 11 expression construct, whereas open bars represent % apoptosis of Jurkat cells that have been transfected with the pEF control vector. Apoptosis was measured by two-color cytometry using mitoTracker Red.

Figure 5B is a bar graph showing the survival of GFP positive cells that have been successfully transfected, 4 days after transfection. The survival of the transfected cells was examined after co-transfection with either the control vector (pEF1), or thc galectin 11 expression vector (pEF-Leg11). Therc wcre about 4 times more surviving GFP positive cells after transfection with pEF1 than with pEF-Leg11.

Detailed Description

The present invention provides isolated nucleic acid molecules comprising a polynucleotide encoding a galectin 11 polypeptide having the amino acid sequence shown in Figure 1 (SEQ ID NO:2), which was determined by sequencing a cloned cDNA. The nucleotide sequence shown in Figure 1 (SEQ ID NO:1) was obtained by sequencing the HJACE54 plasmid which was deposited on May 16, 1997 at the American Type Culture Collection, 10801 University Blvd, Manassas, Va 20110-2209 ~~12301 Park Lawn Drive, Rockville, Maryland 20852~~, and given accession number 209053. The galectin 11 polypeptides of the present invention share sequence homology with rat galectin 5, chicken galectin 3, and human galectin 8 gene products (see, e.g., Figurc 2; SEQ ID NOS: 3-4).

The invention further provides for fragments, variants, derivatives and analogs of galectin 11 polynucleotides and polypeptides encoded thereby, and antibodies which bind these polypeptides.

Definitions

The following definitions are provided to facilitate understanding of certain terms used frequently herein.

"Functional activity" or "biological activity" refers to galectin 11 polypeptides, fragments, derivatives, variants, and analogs, which are functionally active, i.e., they are capable of displaying one or more functional activities associated with a complete or mature galectin 11 polypeptide. Such functional activities include, but are not limited to, biological activity (such as, for example, the ability to bind a β-galactoside sugar, the ability to agglutinate trypsin-treated rabbit erythrocytes and/or to induce apoptosis), antigenicity (ability to bind or compete with a galectin 11 polypeptide for binding to an anti-galectin 11 antibody), immunogenicity (ability to generate antibody which binds to a galectin 11 polypeptide), the ability to form dimers with galectin 11 polypeptides of the invention, and the ability to bind to other galectins and/or a receptor or ligand for galectin 11. Polynucleotides encoding polypeptides having galectin 11 functional or biological activity, and the complementary strand of these polynucleotides are also encompassed by the invention.

"Polynucleotide" generally refers to any polyribonucleotide or polydeoxyribonucleotide, which may be unmodified RNA or DNA, or modified RNA or DNA. "Polynucleotides" include, without limitation single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term polynucleotide also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications have been made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

"Polypeptide" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres. "Polypeptide" refers to both short chains, commonly referred to as peptides,

oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptides" include amino acid sequences modified either by natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be apprcciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993 and Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter et al., "Analysis for protein modifications and nonprotein cofactors", Meth Enzymol (1990) 182:626-646 and Rattan et al., "Protein Synthesis: Posttranslational Modifications and Aging", Ann NY Acad Sci (1992) 663:48-62.)

"Variant" as the term is used herein, is a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains functional or biological activity of galectin 11. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the

polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

"Antibodies" as used herein includes polyclonal and monoclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, including the products of an Fab or other immunoglobulin expression library.

Nucleic Acid Molecules

Unless otherwise indicated, all nucleotide sequences determined by sequencing a DNA molecule herein were determined using an automated DNA sequencer (such as the Model 373 from Applied Biosystems, Inc.), and all amino acid sequences of polypeptides encoded by DNA molecules determined herein were predicted by translation of a DNA sequence determined as above. Therefore, as is known in the art for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art. As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

Using the information provided herein, such as the nucleotide sequence in Figure 1, a nucleic acid molecule of the present invention encoding a galectin 11 polypeptide may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA as starting material. Illustrative of the

invention, the nucleic acid molecule described in Figure 1 (SEQ ID NO:1) was discovered in a cDNA library derived from G1 phase Jurkat T-cells. This gene was also identified in cDNA libraries generated from human neutrophil and human infant adrenal gland. Polynucleotides of the invention can also be obtained from natural sources such as mRNA or genomic DNA using techniques known in the art, or can be chemically synthesized using techniques known in the art.

The human galectin 11 gene is located on chromosome 11 and contains 5 exons (see, e.g., Figure 4). The nucleotide sequence of the galectin 11 cDNA of Figure 1 (SEQ ID NO:1) is 865 nucleotides in length (830 nucleotides discounting the poly A tail of the cDNA) and has an open reading frame encoding a protein of 133 amino acid residues. There is a predicted initiation codon at nucleotides 49-51 of the nucleotide sequence depicted in Figure 1 (SEQ ID NO:1), located on the second exon of the gene. The galectin 11 protein shown in Figure 1 (SEQ ID NO:2) shares homology with the translation product of rat galectin 5, chicken galectin 3, and human galectin 8 (see, e.g., Figure 2). Additionally, as further discussed below, galectin 11 induces apoptosis of transfected T-cells (see Example 5 and Figures 5A and 5B). These findings indicate that galectin 11 functions in a manner similar to other previously characterized galectins and therefore, that galectin 11 is important in the regulation of cell growth disorders, autoimmune diseases, cancer, and inflammatory diseases.

As one of ordinary skill would appreciate, due to the possibilities of sequencing errors discussed above, as well as the variability of processing sites for different known proteins, the predicted galectin 11 polypeptide encoded by the deposited cDNA comprises about 133 amino acid residues, but may be anywhere in the range of 125-150 amino acids.

As indicated, nucleic acid molecules of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced synthetically. The DNA may be double-stranded or single-stranded. Single-stranded DNA or RNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the complementary or anti-sense strand.

By "isolated" nucleic acid molecule(s) is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA molecules of the present invention.

Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically. In a specific embodiment, "isolated" nucleic acid molecules of the invention comprise all or a portion of the coding region of galectin 11, as disclosed in Figure 1 (SEQ ID NO:1). In further embodiments, the isolated nucleic acid molecules of the invention are not associated with (e.g., linked to) polynucleotides encoding the entire coding region, or alternatively, a portion thereof, of any, or 1, 2, 3, 4, 5, 10, 15, 20, 25 or 50 of the genes that flank galectin 11 on chromosome 11.

Isolated nucleic acid molecules of the present invention include DNA molecules comprising an open reading frame (ORF) or a portion of an ORF shown in Figure 1 (SEQ ID NO:1); and DNA molecules which comprise a sequence substantially different from those described above, but which due to the degeneracy of the genetic code, still encode the galectin 11 protein. Of course, the genetic code is well known in the art. Thus, it would be routine for one skilled in the art to generate such degenerate variants.

In specific embodiments, the invention provides isolated nucleic acid molecules encoding the full length galectin 11 polypeptide depicted in Figure 1 (SEQ ID NO:2), and galectin 11 nucleic acid molecules encoding the galectin 11 polypeptide sequence encoded by the cDNA clone contained in the plasmid deposited as ATCC Deposit No. 209053, on May 16, 1997. In a further embodiment, nucleic acid molecules are provided encoding the full length galectin 11 polypeptide lacking the N-terminal methionine. The invention further provides an isolated nucleic acid molecule having the nucleotide sequence shown in Figure 1 (SEQ ID NO:1) or the nucleotide sequence of the galectin 11 cDNA contained in the above-described deposited clone, or a nucleic acid molecule having a sequence complementary to one of the above sequences. Such isolated molecules, particularly DNA molecules, have uses which include, but are not limited to, probes for gene mapping by *in situ* hybridization with chromosomes, and for detecting expression of the galectin 11 gene in human tissue, for instance, by Northem blot analysis.

The present invention is further directed to fragments of the isolated nucleic acid molecules described herein. By a fragment of an isolated nucleic acid molecule having the nucleotide sequence of the deposited cDNA, the nucleotide sequence shown in Figure 1 (SEQ ID NO:1), or the complementary strand thereto, is intended fragments of at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. By a fragment at least 20 nt in length, for example, is intended fragments which include 20 or more contiguous bases from the nucleotide sequence of the deposited cDNA or the nucleotide sequence as shown in Figure 1 (SEQ ID NO:1) or the complementary strand thereto. Also encompassed by the invention are DNA fragments comprising 50, 100,

150, 200, 250, 300, 350, 365, 370, 375, 380, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850 contiguous nucleotides of the sequence shown in Figure 1 (SEQ ID NO:1), the strand complementary thereto, or contained in the deposited clone. The present invention also encompasses fragments corresponding to most, if not all, of the nucleotide sequence of the deposited cDNA or as shown in Figure 1 (SEQ ID NO:1) or the complimentary strand thereto. In further embodiments, the polynucleotide fragments of the invention comprise a sequence which encodes amino acids 1-14, 1-20, 1-40, 1-66, 2-67, 3-8, 3-67, 5-108, 5-128, 10-17, 10-20, 12-16, 13-20, 13-68, 14-67, 23-40, 20-50, 40-108, 41-60, 47-61, 47-108, 47-128, 50-100, 61-80, 65-108, 65-128, 66-108, 76-88, 81-100, 88-108, 88-128, 95-101, 101-133, 108-120, 114-128, and/or 114-128 of the amino acid sequence depicted in Figure 1 (SEQ ID NO:2). In preferred embodiments, polynucleotide fragments of the invention encode a polypeptide which demonstrates a galectin 11 functional activity. Fragments of the invention have numerous uses which include, but are not limited to, diagnostic probes and primers as discussed herein.

Preferred nucleic acid fragments of the present invention include nucleic acid molecules encoding epitope-bearing portions of the galectin 11 protein. In particular, such nucleic acid fragments of the present invention include nucleic acid molecules encoding: a polypeptide comprising amino acid residues from about 65-70 and 118-124 in Figure 1 (SEQ ID NO:2). The inventors have determined that the above polypeptide fragments are antigenic regions of the galectin 11 protein. Methods for determining other such epitope-bearing portions of the galectin 11 protein are described in detail below.

In other embodiments, the invention provides an isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to all or a portion of a polynucleotide (including fragments) described herein, such as, for example, the complementary strand of the DNA sequence depicted in Figure 1 (SEQ ID NO:1), the cDNA clone contained in ATCC Deposit No. 209053, on May 16, 1997, or fragments thereof. By "stringent hybridization conditions" is intended overnight incubation at 42°C in a solution comprising: 50% formamide, 5X SSC ~~(150 mM NaCl, 15mM trisodium citrate)~~ (750 mM NaCl, 75 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5X Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1 X SSC at 65°C.

By a polynucleotide which hybridizes to a portion of a polynucleotide is intended a polynucleotide (either DNA or RNA) hybridizing to at least about 15 nucleotides (nt), and more preferably at least about 20, still more preferably at least

about 30, 50, 60, 75, 100, 150, 175, 200, 250, 300, 350 nt preferable about 30-70nt, or 80-150 nucleotides, or the entire length of the reference polynucleotide. By a portion of a polynucleotide of at least "20 nt in length", for example, is intended 20 or more contiguous nucleotides from the nucleotide sequence of the reference polynucleotide (e.g., the deposited cDNA or the nucleotide sequence as depicted in Figure 1 (SEQ ID NO:1). In specific embodiments, the polynucleotide hybridizes to nucleotides 0-20, 0-25, 0-30, 0-50, 51-100, 80-100, 101-200, 201-300, 301-400, 401-450, 451-500, 501-550, 551-600, 601-700, 701-750, 751-780, and/or 780-820 of the nucleotide sequence disclosed in Figure 1 (SEQ ID NO:1). In other specific embodiments, the polynucleotide hybridizes to a nucleotide sequence which encodes amino acid residues 1-14, 10-20, 20-50, 50-100, 100-133 of the amino acid sequence depicted in Figure 1 (SEQ ID NO:2). These polynucleotides have uses which include, but are not limited to, diagnostic probes and primers, as discussed above and in more detail below.

Of course, a polynucleotide which hybridizes only to a poly A sequence (such as the 3' terminal poly(A) tract of the galectin 11 cDNA shown in Figure 1 (SEQ ID NO:1), or to a complementary stretch of T (or U) residues, would not be included in a polynucleotide of the invention used to hybridize to a portion of a nucleic acid of the invention, since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone generated using an oligo-dT primer).

As indicated, nucleic acid molecules of the present invention which encode a galectin 11 polypeptide may include, but are not limited to, those encoding the amino acid sequence of the polypeptide, by itself; the coding sequence for the polypeptide and additional sequences, such as those encoding an amino acid leader or secretory sequence, such as a pre-, or pro- or prepro- protein sequence; the coding sequence of the polypeptide, with or without the aforementioned additional coding sequences, together with additional, non-coding sequences, including for example, but not limited to, introns and non-coding 5′ and 3′ sequences, such as the transcribed, non-translated sequences that play a role in transcription, mRNA processing, including splicing and polyadenylation signals, for example - ribosome binding and stability of mRNA; an additional coding sequence which codes for additional amino acids, such as those which provide additional functionalities. Thus, the sequence encoding the polypeptide may be fused to a marker sequence, such as a sequence encoding a peptide which facilitates purification of the fused polypeptide. In certain preferred embodiments of this aspect of the invention, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (Qiagen, Inc.), among others, many of which

are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. The "HA" tag is another peptide useful for purification which corresponds to an epitope derived from the influenza hemagglutinin protein, which has been described by Wilson et al., Cell 37:767-778 (1984). As discussed below, other such fusion proteins include the galectin 11 fused to Fc at the N- or C-terminus.

The present invention further relates to variants of the nucleic acid molecules of the present invention, which encode a portion (i.e., fragments), analogs or derivatives of the galectin 11 protein. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. *Genes II*, Lewin, B., ed., John Wiley & Sons, New York (1985). Non-naturally occurring variants may be produced using art-known mutagenesis techniques.

Such variants include those produced by nucleotide substitutions, deletions or additions which may involve one or more nucleotides Particularly preferred are variants in which the nucleotide sequence encoding 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 15, 20, 25, 30, 35, 40, 50, or, 20-15, 15-10, ~~10-5 1-5,, 1-3, or 1-2 amino acids~~ 10-5, 1-5, 1-3 or 1-2 amino acid of a polypeptide of the invention are substituted, deleted, or added in any combination. The variants may be altered in coding regions, non-coding regions, or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Especially preferred among these are silent substitutions, additions and deletion, which do not alter the properties and activities of the galectin 11 protein or portions thereof. Also especially preferred in this regard are conservative substitutions.

Further embodiments of the invention include isolated nucleic acid molecules comprising a polynucleotide having a nucleotide sequence at least 75%, 80%, 85%, or 90% identical, and more preferably at least 95%, 96%, 97%, 98% or 99% or 98-99% identical to (a) a nucleotide encoding amino acids 1 to 133 of SEQ ID NO:2; (b) a nucleotide encoding amino acids 2 to 133 of SEQ ID NO:2; (c) a nucleotide sequence of the galectin 11 polypeptide encoded by the cDNA contained in ATCC Deposit No. 209053; or (d) fragments and other polynucleotide sequences of the invention as described herein.

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence encoding a galectin 11 polypeptide is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five nucleotide mismatches per each 100 nucleotides of the reference nucleotide sequence encoding the

galectin 11 polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5′ or 3′ terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular nucleic acid molecule is at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the nucleotide sequence shown in Figure 1 or to the nucleotides sequence of the deposited cDNA clone can be determined conventionally using known computer programs such as the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711. Bestfit uses the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2: 482-489 (1981), to find the best segment of homology between two sequences. When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference nucleotide sequence and that gaps in homology of up to 5% of the total number of nucleotides in the reference sequence are allowed.

In a specific embodiment, the identity between a reference (query) sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, is determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. 6:237-245 (1990)). Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identify are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter. According to this embodiment, if the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction is made to the results to take into consideration the fact that the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends,

relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. A determination of whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of this embodiment. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score. For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are made for the purposes of this embodiment.

The present application is directed to nucleic acid molecules at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequence shown in Figure 1 (SEQ ID NO:1), nucleic acid sequence of the deposited cDNA clone, and fragments thereof, irrespective of whether they encode a polypeptide having galectin 11 functional activity. This is because even where a particular nucleic acid molecule does not encode a polypeptide having galectin 11 functional activity, one of skill in the art would still know how to use the nucleic acid molecule, for instance, as a hybridization probe or a polymerase chain reaction (PCR) primer. Uses of the nucleic acid molecules of the present invention that do not encode a polypeptide having galectin 11 functional activity include, *inter alia,* (1) isolating the galectin 11 gene or allelic variants thereof in a cDNA library; (2) *in situ* hybridization (e.g., FISH) to metaphase chromosomal spreads to provide precise chromosomal location of the galectin 11 gene, as described in Verma et al., *Human Chromosomes: A Manual of Basic Techniques*, Pergamon

Press, New York (1988); (3) use in linkage analysis as a marker for chromosome 11; and (4) Northern Blot analysis for detecting galectin 11 mRNA expression in specific tissues.

Preferred, however, are nucleic acid molecules having sequences at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequence shown in Figure 1 (SEQ ID NO:1), nucleic acid sequence of the deposited cDNA clone, the nucleic acid encoding the polypeptide shown in Figure 1 (SEQ ID NO:2), and fragments thereof, which do, in fact, encode a polypeptide having galectin 11 functional activity. By a polypeptide having galectin 11 functional activity is intended polypeptides exhibiting activity similar, but not necessarily identical, to a functional activity of the galectin 11 protein of the invention, as measured in a particular assay. For example, galectin 11 protein activity can be measured using a β-galactoside sugar (e.g., thiodigalactoside or lactose) binding assay, an assay for apoptosis and/or an assay for agglutination of trypsin-treated rabbit erythrocytes, as further described below.

Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the nucleic acid molecules having a sequence at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical to the nucleic acid sequence of the deposited cDNA, the nucleic acid sequence shown in Figure 1 (SEQ ID NO:1), the nucleic acid encoding the polypeptide shown in Figure 1 (SEQ ID NO:2), or fragment thereof, will encode "a polypeptide having galectin 11 functional activity". In fact, since numerous degenerate variants of these nucleotide sequences encode the same polypeptide, this will be clear to the skilled artisan even without performing the above described comparison assay. It will be further recognized in the art that, for such nucleic acid molecules that are not degenerate variants, a reasonable number will also encode a polypeptide having galectin 11 activity. This is because the skilled artisan is fully aware of amino acid substitutions that are either less likely or not likely to significantly effect protein function (e.g., replacing one aliphatic amino acid with a second aliphatic amino acid).

For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie et al., Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions, Science 247:1306-1310 (1990), wherein the authors indicate that proteins are surprisingly tolerant of amino acid substitutions.

Vectors and Host Cells

The present invention also relates to vectors which include the isolated DNA molecules of the present invention, host cells which are genetically engineered with the polynucleotides and/or recombinant vectors of the invention, and the production of galectin 11 polypeptides and fragments, variants, derivatives, and analogs thereof, by recombinant techniques.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged *in vitro* using an appropriate packaging cell line and then transduced into host cells.

In one embodiment, the DNA of the invention is operatively associated with an appropriate heterologous regulatory element (e.g., promoter or enhancer), such as the phage lambda PL promoter, the *E. coli lac*, *trp* and *tac* promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters and enhancers will be known to the skilled artisan.

In embodiments in which vectors contain expression constructs, these constructs will further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase or neomycin resistance for eukaryotic cell culture and tetracycline or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptococcus staphylococci*, *Bacillus subtilis*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, HEK293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Selection of appropriate vectors and promoters for expression in a host cell is a well known procedure and the requisite techniques for expression vector construction, introduction of the vector into the host, and expression in the host are routine skills in the art. A great variety of expression vectors can be used to express galectin 11 polypeptides and fragments, variants, derivatives, and analogs of the invention. Such vectors include chromosomal, episomal and virus-derived vectors e.g., vectors derived

from bacterial plasmids, from bacteriophage, from yeast episomes, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids, all may be used for expression in accordance with this aspect of the present invention. Generally, any vector suitable to maintain, propagate or express polynucleotides to express a polypeptide in a host may be used. The appropriate nucleotide sequence may be inserted into an expression vector system by any of a variety of known technique, such as for example, those set forth in Ausubel et al., eds., 1989, Current Protocols in Molecular Biology, Green Publishing Associates, Inc., and John Wiley & Sons, Inc., New York.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

The present invention also relates to host cells containing the vector constructs discussed herein, and additionally encompasses host cells containing nucleotide sequences of the invention that are operably associated with one or more heterologous control regions (e.g., promoters and/or enhancers) using techniques known in the art. As discussed above, the host cell can be a higher eukaryotic cell, such as a mammalian cell (e.g., a human derived cell), or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. The host strain may be chosen which modulates the expression of the inserted gene sequences, or modifies and processes the gene product in the specific fashion desired. Expression from certain promoters can be elevated in the presence of certain inducers; thus expression of the genetically engineered polypeptide may be controlled. Furthermore, different host cells have characteristics and specific mechanisms for the translational and post-translational processing and modification (e.g., glycosylation, phosphorylation, cleavage) of proteins. Appropriate cell lines can be chosen to ensure the desired modifications and processing of the foreign protein expressed.

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the desired polypeptide using techniques

known in the art. These signals may be endogenous to the polypeptide or they may be heterologous signals.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., *Basic Methods In Molecular Biology* (1986).

The polypeptide may be expressed in a modified form, such as a fusion protein (comprising the polypeptide joined via a peptide bond to a heterologous protein sequence (of a different protein)), and may include not only secretion signals, but also additional heterologous functional regions. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell, during purification, or during subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate purification, among others, are familiar and routine techniques in the art. Altematively, such a fusion protein can be made by protein synthetic techniques, e.g., by use of a peptide synthesizer.

A preferred fusion protein comprises a heterologous region from immunoglobulin that is useful to solubilize proteins. For example, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is thoroughly advantageous for use in therapy and diagnosis and thus results, for example, in improved pharmacokinetic properties (EP-A 0232 262). On the other hand, for some uses it would be desirable to be able to delete the Fc part after the fusion protein has been expressed, detected and purified in the advantageous manner described. This is the case when Fc portion proves to be a hindrance to use in therapy and diagnosis, for example when the fusion protein is to be used as antigen for immunizations. In drug discovery, for example, human proteins, such as, hIL5- has been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. See, Bennett et al., J. Md. Recog. 8:52-58 (1995) and Johanson et al., J. Biol. Chem. 270(16):9459-9471 (1995).

The galectin 11 protein can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation,

acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, plant, insect, teleost, avian, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue or may be missing an initial methionine residue, in some cases as a result of host-mediated processes.

Galectin 11 Polypeptides and Fragments

The invention further provides an isolated galectin 11 polypeptide having the amino acid sequence encoded by the deposited cDNA, the amino acid sequence depicted in Figure 1 (amino acid residues 1-133 of SEQ ID NO:2), the amino acid sequence depicted in Figure 1 less the amino terminal methionine (amino acid residues 2-133 of SEQ ID NO:2), polypeptides which are encoded by a polynucleotide that hybridizes under stringent hybridization conditions to a polynucleotide sequence encoding a polypeptide having the amino acid sequence depicted in Figure 1 (SEQ ID NO:2) and/or containcd in the deposited clone, and fragments, variants, derivatives and analogs of these polypeptides.

The polypeptides of the present invention are preferably provided in an isolated form. By "isolated polypeptide" is intended a polypeptide removed from its native environment. Thus, a polypeptide produced and/or contained within a recombinant host cell is considered isolated for purposes of the present invention. Also intended as an "isolated polypeptide" are polypeptides that have been purified, partially or substantially, from a recombinant host cell. For example, a recombinantly produced version of the galectin 11 polypeptide can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988).

It will be recognized in the art that some amino acid sequences of the galectin 11 polypeptide can be varied without significant effect of the structure or function of the protein. If such differences in sequence are contemplated, it should be remembered that there will be critical areas on the protein which determinc activity.

Thus, the invention further includes variations of the galectin 11 polypeptide which show substantial galectin 11 polypeptide functional activity or which include regions of galectin 11 protein such as the protein portions discussed below. Such mutants include deletions, insertions, inversions, repeats, and type substitutions.

As indicated above, guidance concerning which amino acid changes are likely to be phenotypically silent can be found in Bowie et al., Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions, Science 247:1306-1310 (1990). Thus, a fragment, variant, derivative, or analog of the polypeptide of Figure 1 (SEQ ID NO:2), or that encoded by the deposited cDNA, include (i) one in which at least one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue(s), and more preferably at least one but less than ten conserved amino acid residues) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature polypeptide, such as an IgG Fc fusion region peptide or leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments, variants, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

Of particular interest are substitutions of one or more charged amino acids with other charged amino acid and with neutral or negatively charged amino acids. The latter results in proteins with reduced positive charge to improve the characteristics of the galectin 11 protein. The prevention of aggregation is highly desirable. Aggregation of proteins not only results in a loss of activity but can also be problematic when preparing pharmaceutical formulations, because they can be immunogenic.(Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36:838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993)).

As indicated, changes are preferably of a minor nature, such as conservative amino acid substitutions that do not significantly affect the folding or activity of the protein (see Table 1).

TABLE 1. Conservative Amino Acid Substitutions

Aromatic	Phenylalanine
	Tryptophan
	Tyrosine
Hydrophobic	Leucine
	Isoleucine
	Valine
Polar	Glutamine
	Asparagine
Basic	Arginine
	Lysine
	Histidine
Acidic	Aspartic Acid
	Glutamic Acid
Small	Alanine
	Serine
	Threonine
	Methionine
	Glycine

In the specific embodiments, the number of additions, substitutions and/or deletions in the amino acid sequence of Figure 1 (SEQ ID NO:2) and/or any of the polypeptide fragments described herein is 50, 40, 35, 30, 25, 20, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1, or 15-20, 15-10, 5-10, 1-5, 1-3, or 1-2.

Amino acid residues of the galectin 11 polypeptide, fragment, variant, derivative, or analog of the present invention that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, Science 244:1081-1085 (1989)). The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity or functional activity, such as, receptor binding, β-galactoside (e.g., thiodigalactoside or lactose)

binding, the ability to agglutinate trypsin-treated rabbit erythrocytes, or the ability *in vitro* or *in vivo* to induce apoptosis. Sites that are critical for ligand-receptor binding can also be determined by structural analysis such as crystallization, nuclear magnetic resonance or photoaffinity labeling (Smith et al., J. Mol. Biol. 224:899-904 (1992) and de Vos et al., Science 255:306-312 (1992)).

The present invention also encompasses polypeptides which are at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%, or, 97-99% identical to the polypeptides described above. By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a reference amino acid sequence of a galectin 11 polypeptide is intended that the amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the reference amino acid of the galectin 11 polypeptide. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequence shown in Figure 1 (SEQ ID NO:2), the amino acid sequence encoded by deposited cDNA clone, or fragments thereof, can be determined conventionally using known computer programs such the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711. When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference amino acid sequence and that gaps in homology of up to 5% of the total number of amino acid residues in the reference sequence are allowed.

In another embodiment, the identity between a reference (query) sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, is determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. 6:237-245 (1990)). Preferred

parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter. According to this embodiment, if the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction is made to the results to take into consideration the fact that the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. A determination of whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of this embodiment. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence. For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C- termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are made for the purposes of this embodiment.

The polypeptides of the present invention have uses which include, but are not limited to, molecular weight marker on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art.

The present invention also encompasses fragments of the above-described polypeptides of the invention. In specific embodiments, these fragments are at least 20, 25, 30, 40, 50, 75, 90, 100, 110, 120, 125, or 130 amino acid residues in length.

For many proteins, including the mature form(s) of a secreted protein, it is known in the art that one or more amino acids may be deleted from the N-terminus or C-terminus without substantial loss of biological function. However, even if deletion of one or more amino acids from the N-terminus or C-terminus of a protein results in modification or loss of one or more biological functions of the protein, other galectin 11 functional activities may still be retained. For example, in many instances, the ability of the shortened protein to induce and/or bind to antibodies which recognize galectin 11 (preferably antibodies that bind specifically to galectin 11) (e.g., that do not cross-react with other previously characterized galectins) will be retained irrespective of the size or location of the deletion. Whether a particular polypeptide lacking N-terminal and/or C-terminal residues of a complete protein retains such immunologic activities can readily be determined by routine methods described herein and otherwise known in the art.

In one embodiment, the present invention further provides polypeptides having one or more residues deleted from the amino terminus of the amino acid sequence of the galectin 11 polypeptide depicted in Figure 1 (SEQ ID NO:2) or encoded by the cDNA of the deposited clone. Particularly, in one embodiment, N-terminal deletions of the galectin 11 polypeptide can be described by the general formula m to 133, where m is a integer from 1 to 132 corresponding to the position of amino acid residue identified in SEQ ID NO:2 and preferably, corresponds to one of the N-terminal amino acid residues identified in the N-terminal deletions specified herein. In specific embodiments, N-terminal deletions of the galectin 11 polypeptide of the invention comprise, or alternatively consist of, amino acid residues: S-2 to S-133; P-3 to S-133; R-4 to S-133; L-5 to S-133; E-6 to S-133; V-7 to S-133; P-8 to S-133; C-9 to S-133; S-10 to S-133; H-11 to S-133; A-12 to S-133; L-13 to S-133; P-14 to S-133; Q-15 to S-133; G-16 to S-133; L-17 to S-133; S-18 to S-133; P-19 to S-133; G-20 to S-133; Q-21 to S-133; V-22 to S-133; I-23 to S-133; I-24 to S-133; V-25 to S-133; R-26 to S-133; G-27 to S-133; L-28 to S-133; V-29 to S-133; L-30 to S-133; Q-31 to S-133; E-32 to S-133; P-33 to S-133; K-34 to S-133; H-35 to S-133; F-36 to S-133; T-37 to S-133; V-38 to S-133; S-39 to S-133; L-40 to S-133; R-41 to S-133; D-42 to S-133; Q-43 to S-133; A-44 to S-133; A-45 to S-133; H-46 to S-133; A-47 to S-133; P-48 to S-133; V-49 to S-133; T-50 to S-133; L-51 to S-133; R-52 to S-133; A-53 to S-133; S-54

to S-133; F-55 to S-133; A-56 to S-133; D-57 to S-133; R-58 to S-133; T-59 to S-133; L-60 to S-133; A-61 to S-133; W-62 to S-133; I-63 to S-133; S-64 to S-133; R-65 to S-133; W-66 to S-133; G-67 to S-133; Q-68 to S-133; K-69 to S-133; K-70 to S-133; L-71 to S-133; I-72 to S-133; S-73 to S-133; A-74 to S-133; P-75 to S-133; F-76 to S-133; L-77 to S-133; F-78 to S-133; Y-79 to S-133; P-80 to S-133; Q-81 to S-133; R-82 to S-133; F-83 to S-133; F-84 to S-133; E-85 to S-133; V-86 to S-133; L-87 to S-133; L-88 to S-133; L-89 to S-133; F-90 to S-133; Q-91 to S-133; E-92 to S-133; G-93 to S-133; G-94 to S-133; L-95 to S-133; K-96 to S-133; L-97 to S-133; A-98 to S-133; L-99 to S-133; N-100 to S-133; G-101 to S-133; Q-102 to S-133; G-103 to S-133; L-104 to S-133; G-105 to S-133; A-106 to S-133; T-107 to S-133; S-108 to S-133; M-109 to S-133; N-110 to S-133; Q-111 to S-133; Q-112 to S-133; A-113 to S-133; L-114 to S-133; E-115 to S-133; Q-116 to S-133; L-117 to S-133; R-118 to S-133; E-119 to S-133; L-120 to S-133; R-121 to S-133; I-122 to S-133; S-123 to S-133; G-124 to S-133; S-125 to S-133; V-126 to S-133; Q-127 to S-133; L-128 to S-133; of SEQ ID NO:2. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Further embodiments of the invention are directed to C-terminal deletions of the galectin 11 polypeptide described by the general formula 1 to n, where n is an integer from 2 to 132 corresponding to the position of amino acid residue identified in SEQ ID NO:2 and preferably, corresponds to one of the C-terminal amino acid residues identified in the C-terminal deletions specified herein. In specific embodiments, C terminal deletions of the galectin 11 polypeptide of the invention comprise, or alternatively, consist of, amino acid residues: M-1 to H-132; M-1 to V-131; M-1 to C-130; M-1 to Y-129; M-1 to L-128; M-1 to Q-127; M-1 to V-126; M-1 to S-125; M-1 to G-124; M-1 to S-123; M-1 to I-122; M-1 to R-121; M-1 to L-120; M-1 to E-119; M-1 to R-118; M-1 to L-117; M-1 to Q-116; M-1 to E-115; M-1 to L-114; M-1 to A-113; M-1 to Q-112; M-1 to Q-111; M-1 to N-110; M-1 to M-109; M-1 to S-108; M-1 to T-107; M-1 to A-106; M-1 to G-105; M-1 to L-104; M-1 to G-103; M-1 to Q-102; M-1 to G-101; M-1 to N-100; M-1 to L-99; M-1 to A-98; M-1 to L-97; M-1 to K-96; M-1 to L-95; M-1 to G-94; M-1 to G-93; M-1 to E-92; M-1 to Q-91; M-1 to F-90; M-1 to L-89; M-1 to L-88; M-1 to L-87; M-1 to V-86; M-1 to E-85; M-1 to F-84; M-1 to F-83; M-1 to R-82; M-1 to Q-81; M-1 to P-80; M-1 to Y-79; M-1 to F-78; M-1 to L-77; M-1 to F-76; M-1 to P-75; M-1 to A-74; M-1 to S-73; M-1 to I-72; M-1 to L-71; M-1 to K-70; M-1 to K-69; M-1 to Q-68; M-1 to G-67; M-1 to W-66; M-1 to R-65; M-1 to S-64; M-1 to I-63; M-1 to W-62; M-1 to A-61; M-1 to L-60; M-1 to T-59; M-1 to R-58; M-1 to D-57; M-1 to A-56; M-1 to F-55; M-1 to S-54; M-1 to A-53; M-1 to R-52; M-1 to L-51; M-1 to T-50; M-1 to V-49; M-1 to P-48; M-1 to A-47; M-1 to H-46; M-1 to A-45; M-1

to A-44; M-1 to Q-43; M-1 to D-42; M-1 to R-41; M-1 to L-40; M-1 to S-39; M-1 to V-38; M-1 to T-37; M-1 to F-36; M-1 to H-35; M-1 to K-34; M-1 to P-33; M-1 to E-32; M-1 to Q-31; M-1 to L-30; M-1 to V-29; M-1 to L-28; M-1 to G-27; M-1 to R-26; M-1 to V-25; M-1 to I-24; M-1 to I-23; M-1 to V-22; M-1 to Q-21; M-1 to G-20; M-1 to P-19; M-1 to S-18; M-1 to L-17; M-1 to G-16; M-1 to Q-15; M-1 to P-14; M-1 to L-13; M-1 to A-12; M-1 to H-11; M-1 to S-10; M-1 to C-9; M-1 to P-8; M-1 to V-7; M-1 to E-6; of SEQ ID NO:2. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Further embodiments of the invention are directed to polypeptide fragments comprising, or alternatively, consisting of, amino acids described by the general formula m to n, where m and n correspond to any one of the amino acid residues specified above for these symbols, respectively. Polynucleotides encoding these polypeptides are also encompassed by the invention.

In preferred embodiments, the polypeptides of the invention comprise, or alternatively, consist of, amino acid residues: M-1 to L-40; M-1 to W-66; P-3 to L-40; L-5 to L-40; L-5 to S-108; L-5 to L-128; P-3 to L-128; L-5 to L-128; L-5 to G-124; C-9 to C-130; L-13 to L-40; P-14 to L-40; L-40 to S-108; A-47 to S-108; A-47 to L-128; R-65 to S-108; R-65 to L-128; L-88 to S-108; L-88 to L-128; S-108 to L-120; or L-114 to L-128 of SEQ ID NO:2. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Among the especially preferred fragments of the invention are fragments characterized by structural or functional attributes of galectin 11. Such fragments include amino acid residues that comprise alpha-helix and alpha-helix forming regions ("alpha-regions"), beta-sheet and beta-sheet-forming regions ("beta-regions"), turn and turn-forming regions ("turn-regions"). coil and coil-forming regions ("coil-regions"), hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, surface forming regions, and high antigenic index regions (i.e., having an antigenic region of three or more contiguous amino acid residues each of which having an antigenic index of greater than or equal to 1.5) of galectin 11. Certain preferred regions are those set out in Figure 3, and include, but are not limited to, regions of the aforementioned types identified by analysis of the amino acid sequence depicted in Figure 1 (SEQ ID NO:2) using the default parameters of the identified computer programs, such preferred regions include; Garnier-Robson predicted alpha-regions, beta-regions, turn-regions, and coil-regions; Chou-Fasman predicted alpha-regions, beta-regions, turn-regions, and coil-regions; Kyte-Doolittle predicted hydrophilic and hydrophobic regions; Eisenberg alpha and beta amphipathic regions; Emini surface-

forming regions; and Jameson-Wolf high antigenic index regions. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Figure 2 provides a comparison of the galectin 11 polypeptide with other galectins. Identical amino acids shared between the galectins are shaded, while conservative amino acid changes are boxed. By examining the regions of amino acids shaded and/or boxed, the skilled artisan can readily identify conserved domains between the two polypeptides. The amino acid sequences falling within these conserved, shaded and/or boxed domains are contained in the preferred polypeptide fragments of the invention.

Representative examples of polypeptide residue fragments of the invention including, for example, fragments from about amino acid number 1-20, 1-66, 5-108, 5-128, 21-40, 40- 108, 41-60, 47-108, 47-128, 61-80, 65-108, 65-128, 81-100, 88-108, 88-128, 108-120; 114-128; and 101 to the end of the galectin 11 polypeptide depicted in Figure 1 (SEQ ID NO:2). In this context, "about" includes the particularly recited ranges larger or smaller by several , 5, 4, 3, 2, or 1 amino acid at either end or at both extremes.

In another embodiment, the invention provides a peptide or polypeptide comprising an epitope-bearing portion of a polypeptide of the invention. The epitope of this polypeptide portion is an immunogenic or antigenic epitope described herein. An "immunogenic" epitope is defined as a part of a protein that elicits an antibody response when the whole protein is the immunogen. On thc other hand, a region of a protein molecule to which an antibody can bind is defined as an antigenic epitope. The number of immunogenic epitopes of a protein generally is less than the number of antigenic epitopes. See, for instance, Geysen ct al., Proc. Natl. Acad. Sci. USA 81:3998-4002 (1983). As to the selection of peptides or polypeptides bearing an antigenic epitope (i.e., that contain a region of a protein molecule to which an antibody can bind), it is well known in that art that relatively short synthetic peptides that mimic part of a protein sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked protein. See, for instance, Sutcliffe et al., Scicnce 219:660-666 (1993). Peptides capable of eliciting protein-reactive sera are frequently represented in the primary sequence of a protein, can be characterized by a set of simple chemical rules, and are confined neither to immunodominant regions of intact proteins (i.e., immunogenic epitopes) nor to the amino or carboxyl terminals.

Antigenic epitope-bearing peptides and polypeptides of the invention are therefore useful to raise antibodies, including monoclonal antibodies, that bind specifically to a polypeptide of the invention. See, for instance, Wilson et al., Cell 37:767-778 (1984).

Antigenic epitope-bearing peptides and polypeptides of the invention preferably contain a sequence of at least 7, 9, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 110 or 120 contiguous amino acid residues of the amino acid sequence depicted in Figure 1 (SEQ ID NO:2).

Non-limiting examples of antigenic polypeptides or peptides that can be used to generate galectin 11-specific antibodies include: a polypeptide comprising amino acid residues from about 65-70 and 118-124 in Figure 1 (SEQ ID NO:2). As indicated above, the inventors have determined that the above polypeptide fragments are antigenic regions of the galectin 11 protein.

The epitope-bearing peptides and polypeptides of the invention may be produced by any conventional means. See generally, Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985). This Simultaneous Multiple Peptide Synthesis (SMPS) process is further described in U.S. Patent No. 4,631,211 to Houghten et al. (1986).

As one of skill in the art will appreciate, galectin 11 polypeptides of the present invention such as, for example, epitope-bearing fragments of galectin 11, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life *in vivo*. This has been shown, e.g., for fusion proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins (EPA 394, 827; Traunecker et al., Nature 331:84-86 (1988)). Fusion proteins that have a disulfide-linked dimeric structure due to the IgG part can also be more efficient in binding and neutralizing other molecules than the monomeric galectin 11 protein or protein fragment alone (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995)).

Polypeptides of the invention include polypeptides encoded by polynucleotides that hybridize (e.g., under stringent hybridization conditions) to the polynucleotide sequence depicted in Figure 1 (SEQ ID NO:1), the complementary strand thereto, and/or the nucleotide sequence contained in the deposited clone. In specific embodiments, the polypeptides of the invention have galectin 11 functional and/or biological activity.

Assays for Galectin 11 Functional Activity

The functional and/or biological activity of galectin 11 polypeptides, fragments, variants, derivatives and analogs of the invention can be assayed by various methods.

For example, in one embodiment, where one is assaying for the ability to bind or compete with galectin 11 polypeptide for binding to anti-galectin 11 antibody,

various immunoassays known in the art can be used, including, but not limited to, competitive and non-competitive assay systems using techniques such as radioimmunoassays, ELISA, "sandwich" immunoassays, immunoradiometric assays, and diffusion precipitin reactions, immunodiffusion assays, *in situ* immunoassays (using colloidal gold, enzyme or radioisotope labels, for example), western blots, precipitation reactions, agglutination assays (e.g., gel agglutination assays, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays, and immunoelectrophoresis assays, etc. Many means are known in the art for detecting binding in an immunoassay and are within the scope of the present invention.

In another embodiment, the ability of galectin 11 polypeptides, fragments, variants, derivatives and analogs of the invention to bind β-galactoside sugars may be determined using, or routinely modifying, assays known in the art. For example, lactose binding activity of the expressed galectin 11 polypeptides and fragments, variants, derivatives, or analogs thereof, may be assayed by immunodectection of *in situ* binding activity to asialofetuin (Sigma) immobilized on nitrocellulose (Amersham) (Madsen et al., J. Biol. Chem. 270(11):5823-5829 (1995)). For example, in one assay, thirty µg of asialofetuin dissolved in 3 µl of water is spotted on a 1-cm^2 strip of nitrocellulose. The nitrocellulose pieces are then placed in a 24-well tissue culture plant and incubated overnight in buffer B (58 mM NA_2HPO_4, 18 mM KH_2PO_4, 75 mM NaCl, 2 mM EDTA, and 3% BSA, pH 7.2) with constant agitation at 22°C. Following incubation, the blocking medium is aspirated and the nitrocellulose pieces are washed three times in buffer A (58 mM Na_2HPO_4 18 mM KH_2 PO_4, 75 mM NaCl, 2 mM EDTA, 4 mM β-mercaptoethanol and 0.2% BSA, pH 7.2). Cell extracts (preferably, COS cells) are prepared containing 1% BSA and either with or without 150 mM lactose (105 µl of primary extract, 15 µl of 10% BSA in buffer A and either 30 µl of 0.75 M lactose in buffer A or 30 µl of buffer A). The immobilized asialofetuin is incubated with the extracts for 2 h and washed 5 times in buffer A. The nitrocellulose pieces are then fixed in 2% formalin in PBS (58 mM Na_2HPO_4, 18 mM KH_2PO_4, 75 mM NaCl, 2 mM EDTA pH 7.2) for 1 h to prevent loss of bound galectin 11. Following extensive washing in PBS the pieces are incubated with a rabbit anti-galectin 11. Polyclonal serum (generated using techniques known in the art) diluted 1:100 in PBS for 2 h at 22°C. The pieces are then washed in PBS and incubated with peroxidase-labeled goat anti-rabbit antibodies (DAKO). Following incubation for 2 h at 22°C, the pieces are washed in PBS and the substrate is added. Nitrocellulose pieces are incubated until the color developed and the reaction is stopped by washing in distilled water.

The ability of galectin 11 polypeptides, fragments, variants, derivatives and analogs of the invention to agglutinate trypsin-treated rabbit erythrocytes can routinely be assayed using techniques known in the art.

The ability of the galectin 11 polypeptides, fragments, variants, derivatives and analogs of the invention to induce apoptosis of T-cells may be determined using, or routinely modifying, techniques described herein (see e.g., Example 5) or otherwise known in the art. See e.g., Perillo et al., Nature 378:736-739 (1995); Chinnaiyan et al., Cell 81:505-512 (1995); Boldin et al., J. Biol. Chem. 270:7795-7798 (1995); Kischkel et al., EMBO J. 14:5579-5585 (1995); Chinnaiyan et al., J. Biol. Chem. 271:4961-4965 (1996); the contents of each of which is herein incorporated by reference in its entirety).

The galectin 11 polynucleotides and polypeptides, and fragments, variants derivatives and analogs thereof; and antibodies, agonists and antagonists thereto; can be tested *in vivo* for the desired therapeutic or prophylactic activity. For example, such compounds can be tested in suitable animal model systems prior to testing in humans. Such animal models include, but are not limited to, rats, mice, chickens, cows, monkeys, rabbits, etc. Such testing may also be utilized to routinely determine dosage for delivery to human patients. For *in vivo* testing prior to administration to human, any animal model system known in the art may be used (see, for example, Levi et al., Eur. J. Imun. 13:500-507 (1983); and Offner et al., J. Neuroimmunol 28:177-184 (1990)). For example, an animal model useful for the study of the treatment of human MS is experimental allergic encephalomyelitis (EAE). EAE is an experimentally induced disease that shares many of the same clinical and pathological symptoms of MS (Martin et al., Ann. Rev. Immunol. 10:153-187 (1992); Hafler et al., Immunology Today 10:104-107 (1989)). Several studies in rodents have shown that, similar to MS, $CD4^+$ T cells participate in the pathophysiology of EAE, Traugott et al., Cellular Immunology 91:240-254 (1985); Ben-Nun, Eur. J. Immunol. 11:195-199 (1981); Pettinel et al., J. Immunol. 127:1420-1423 (1981). EAE can be induced in certain strains of mice by immunization with myelin in an adjuvant. The immunization activates $CD4^+$ T cells specific for myelin basic protein (MBP) and proteolipid (PLP), Bernard et al., J. Immunol. 114:1537-1540 (1975); Chou et al., J. Immunol. 130:2183-2186 (1983); Kurchroo et al., J. Immunol. 148:3776-3782 (1992). The activated T cells enter the central nervous system and their local action causes both the anatomic pathology and clinical signs, e.g., ascending hind limb paresis leading to paralysis, of the disease. As discussed above, the galectin 1 has been demonstrated to suppress clinical and histological signs of experimental autoimmune encephalomyelitis in rats (Offner et al., J. Neuroimmunol. 28:177-184 (1990)).

Another model system that may bc utilizcd to both study the role of the polypeptides, variants, derivatives and analogs of the invention as a suppresser of immune responses, and to determine effective dosages for doing so, is experimental autoimmune myasthenia gravis (EAMG) in rabbits. EAMG is an autoimmune discase induced by immunization with the purified acctylcholine receptor protein (AChR) and is considered to be a good model for the human disease myasthenia gravis. As further discussed above, galectin 1 has been demonstrated to have a prophylactic and therapeutic action on experimental autoimmune myasthcnia gravis in rabbits (Levi et al., Eur. J. Immunol. 13:500-507 (1983)).

Other art known model assays that may be used to determine the desired therapeutic or prophylactic activity of compounds of the invention (e.g., as a suppresser of immune responses) include, but are not limited to, T-cell proliferation in mixed lymphocyte reaction assays (an art-accepted model for allogencic graft rejection), and murine allograft models known in the art.

Assays described herein or otherwisc known in the art may be applied to routinely determine which galectin 11 polypeptides, fragments, variants, derivatives and analogs of the invention dcmonstrate galectin 11 functional activity and the optimal concentration at which these compounds demonstrate this activity. These assays may additionally be utilized to identify molecules which enhance (agonists) or suppress (antagonists) galectin 11 functional activity.

The exact formulation, route of administration and dosage of the compounds of the invention to be administrated can be choscn by the individual physician in view of the patient's condition (see e.g., Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics." Ch. 1 p. 1).

Other methods will be known to the skilled artisan and are within the scope of the invention.

Antibodies

The polypeptides of the invention and their fragments, variants, derivatives or analogs, or cells expressing them, can also be uscd as immunogens to produce antibodies immunospecific for galectin 11 polypeptide of the invention. The term "immunospecific" means that the antibodies have substantially grcatcr affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

The term "antibody" (Ab) or "monoclonal antibody" (mAb) as used herein is meant to include intact molecules as well as fragments thereof (such as, for example, Fab, and $F(ab')_2$ fragments) which are capable of binding an antigen. Fab, Fab', and F

$(ab')_2$ fragments lack the Fc fragmcnt of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding of an intact antibody (Wahl et al., J. Nucl. Med. 24:316-325 (1983)).

Antibodies according to the present invention may be prepared by any of a variety of standard methods using galectin 11 immunogens of the present invention. For example, antibodies generated against full-length galectin 11 polypcptides can be obtained by administering the polypeptides or epitope-bearing fragments, variants, derivatives, analogs, or cells, to an animal, preferably a nonhuman, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler et al., Nature 256:495-497 (1975)), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., Immunology Today 4:72 (1983)) and the EBV-hybridoma technique (Cole et al., MONOCLONAL ANTIBODIES AND CANCER THERAPY, pp. 77-96, Alan R. Liss, Inc., 1985). Additionally, techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can also be adaptcd to producc single chain antibodies to polypeptides of the invention. Also, transgenic mice, or other organisms including other mammals, may be used to express humanized antibodies.

Antibodies of the invention can be used in methods known in the art relating to the localization and activity of the polypeptide sequences of the invention, e.g., for imaging these polypeptides, measuring levels thereof in appropriate physiological samples, etc. The antibodies also have use in immunoassays and in therapeutics as agonists and antagonists of galectin 11. Additionally, the antibodies of the invention may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography.

Diagnosis and Prognosis

It is believed that certain tissues in mammals with certain diseases (e.g., autoimmune diseases which include, but are not limited to, lupus erythematosus (SLE), rheumatoid arthritis (RA), insulin-dependent diabetes, multiple sclerosis (MS), giant cell arteritis, polyarteritis nodosa, myasthenia gravis, scleroderma, and graft versus host diseasc; graft rejection; mammalian cancers which include, but are not limited, to, mclanoma, renal astrocytoma, Hodgkin's disease, breast, ovarian, prostate, bone, liver, lung, pancreatic, and spleenic cancers; inflammatory diseases; asthma; and allergeic diseases) express significantly altered (e.g., enhanced or decreased) levels of the galectin 11 polypeptide and mRNA encoding the galectin 11 polypeptide when compared to a corresponding "standard" mammal, i.e., a mammal of the same species

not having the disease. Further, it is believed that altered levels of the galectin 11 polypeptide can be detected in certain body fluids (e.g., sera, plasma, urine, and spinal fluid) from mammals with the disorder when compared to sera from mammals of the same species not having the disorder. Thus, the invention provides a diagnostic method useful during diagnosis, which involves assaying the expression level of the gene encoding the galectin 11 polypeptide in mammalian cells or body fluid and comparing the gene expression level with a standard galectin 11 gene expression level, whereby an increase or decrease in the gene expression level over the standard is indicative of the disease.

Where a diagnosis has already been made according to conventional methods, the present invention is useful as a prognostic indicator, whereby patients exhibiting altered galectin 11 gene expression will experience a worse clinical outcome relative to patients expressing the gene at a normal level.

By "assaying" the expression level of the gene encoding the galectin 11 polypeptide" is intended qualitatively or quantitatively measuring or estimating the level of the galectin 11 polypeptide or the level of the mRNA encoding the galectin 11 polypeptide in a first biological sample either directly (e.g., by determining or estimating absolute polypeptide or mRNA level) or relatively (e.g., by comparing to the galectin 11 polypeptide level or mRNA level in a second biological sample).

Nucleic acids for diagnosis may be obtained from a biological sample of a subject, such as from blood, urine, saliva, tissue biopsy or autopsy material, using techniques known in the art. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled galectin 11 nucleotide sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing (see, e.g., Myers et al., Science 230:1242 (1985)). Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (see Cotton et al., Proc. Natl. Acad. Sci. USA 85:4397-4401 (1985)). In another embodiment, an array of oligonucleotides probes comprising galectin 11 polynucleotide sequences or fragments thereof, can be constructed to conduct efficient screening of e.g., genetic mutations. Array technology methods are well known and have general applicability

and can be used to address a varicty of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see for example, Chec et al., Science 274:610-613 (1996)).

The diagnostic assays offcr a process for diagnosing or determining a susceptibility to specific diseases through detection of mutation in the galectin 11 gene by the methods described.

In addition, specific diseases can be diagnosed by methods comprising determining from a sample derivcd from a subject an abnormally decreased or increased level of galectin 11 polypeptide or mRNA. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art, which include, but are not limited to, Northern blot analysis, (Harada et al., Cell 63:303-312 (1990), S1 nuclease mapping (Fijita et al., Cell 49:357-367 (1987)), RNAse protection, the polymerase chain reaction (PCR), reverse transcription in combination with the polymerase chain reaction (RT-PCR) (Makino et al., Technique 2:295-301 (1990), reverse transcription in combination with the ligase chain reaction (RT-LCR) and other hybridization methods.

Assaying galectin 11 polypeptide levels in a biological sample can be by any techniques known in the art, which include, but arc not limited to, radioimmunoassays, competitive-binding assays, Western Blot analysis and enzyme linked immunosorbent assays (ELISAs) and other antibody-based techniques. For example, galcctin 11 polypeptide expression in tissues can be studied with classical immunohistological methods (Jalkanen et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen et al., J. Cell. Biol. 105:3087-3096 (1987)).

Suitable labels are known in the art and include enzyme labels, such as, Glucose oxidase, and radioisotopes, such as iodine (^{125}I, ^{121}I), carbon (^{14}C), sulfur (^{35}S), tritium (^{3}H), indium (^{112}In), and technetium (^{99m}Tc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

Thus in another aspect, the present invention relates to a diagnostic kit for a disease or susceptibility to a disease which comprises:

(a) a galectin 11 polynucleotide, preferably the nucleotide sequence of SEQ ID NO:1, or a fragment thereof ;

(b) a nucleotide sequence complcmcntary to that of (a);

(c) a galectin 11 polypeptide of the invention, preferably the polypeptide of SEQ ID NO:2, or a fragment thereof; or

(d) an antibody to a galectin 11 polypeptide of the invention, preferably to the polypeptide of SEQ ID NO: 2.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

Screening Assays for galectin 11 Agonists or Antagonists

Aberrancies in galectin 11 expression are responsible for many biological functions, including many pathologies. Accordingly, it is desirous to find compounds and drugs which enhance galectin 11 activity or, alternatively, suppress galectin 11 activity. The invention also provides a method of screening compounds to identify those which enhance or suppress galectin 11 activity. An agonist is a compound which increases the natural biological functions of galectin 11 or which functions in a manner similar to galectin 11, while antagonists decrease or eliminate such functions.

Thus, embodiments of the invention are directed to assays designed to identify compounds that interact with (e.g., bind to) galectin 11 polypeptides of the invention, compounds that interfere or enhance the interaction of galectin 11 with its cognate ligands, and to compounds which modulate the galectin 11 gene (i.e., modulate the level of galectin 11 gene expression) or modulate the level of galectin 11 functional or biological activity. Assays may also be used to identify compounds which bind galectin 11 gene regulatory sequences (e.g., promoter sequences) and which may modulate galectin 11 gene expression. See e.g., Platt, J. Biol. Chem. 269:28558-28562 (1994), which is incorporated herein by reference in its entirety.

Thus, polypeptides of the invention may be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991). Further examples of compounds that may be screened include, but are not limited to, peptides such as, for example soluble peptides, including but not limited to, those found: in random peptide libraries (see, e.g., Lam et al., Nature 354:84-86 (1991)), and combinatorial chemistry-derived molecular libraries made of D- and L- configuration amino acid; phosphopeptides (including, but not limited to, members of random or partially degenerate, directed phosphopeptide libraries (see e.g., Songyang et al., Cell 72:767-778 (1993)); antibodies (including but not limited to, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and Fab, F(ab')2, and FAB expression library fragments, and epitope-binding fragments thereof); and small organic or inorganic molecules.

Numerous experimental methods may be used to select and detect compounds that bind galectin 11 polypeptides of the invention and thereby modulate galectin 11 expression or activity, including, but not limited to, protein affinity chromotography, affinity blotting, immunoprecipitation, cross-linking, and library based methods such as protein probing, phage display, the two-hybrid system (Fields and Song, Nature 340:245-246 (1989)), and modified versions of the two-hybrid system (Gyuris et al., Cell 75:791-803 (1993); Zervos et al., Cell 72:223-232 (1993)). See generally, Phizicky et al., Microbiol. Rev. 59:94-123 (1995).

The principle behind assays that identify compounds that bind to galectin 11 polypeptides of the invention involves preparing a reaction mixture of galectin 11 polypeptide and test compound under conditions that allow the two components to interact and bind, thus forming a complex which can be detected in the reaction mixture and purified using techniques known in the art. Accordingly, the assays may simply test binding of a candidate compound to galectin 11.

Further, the assays may simply comprise the steps of combining a candidate compound with a solution containing a galectin 11 polypeptide to form a mixture, and determining the ability of galectin 11 contained in this mixture to bind galectin 11 cognate ligands (e.g., compounds containing a β galactoside sugar and/or molecules expressed on the surface of T-cells), to agglutinate trypsin-treated rabbit erythrocytes, or to induce apoptosis of T-cells, and comparing this ability with that observed for the galectin 11 polypeptide in the same or similar solution under the same or similar conditions, but absent the candidate compound. The ability of the candidate molecule to interfere with binding of galectin 11 to the cognate ligand is reflected in decreased binding of the labeled galectin 11 to the cognate ligand relative to that in the absence of candidate molecule. Molecules which interfere with the ability of galectin 11 to elicit cellular responses (e.g., apoptosis) resulting from galectin 11 binding to its cognate ligand are antagonists. Molecules that enhance galectin 11 induced cellular responses when mixed with galectin-11, or which are able to induce a similar cellular response in the absence of galectin 11, are agonists.

The galectin 11 polynucleotides, polypeptides, and antibodies of the invention may also be used to configure assays for detecting the effect of added compounds on the production of galectin 11 mRNA and protein in cells. For example, an ELISA may be constructed for measuring secreted or cell associated levels of galectin 11 protein using monoclonal and polyclonal antibodies by standard methods known in the art, and this can be used to discover agents which may inhibit or enhance the production of galectin 11 (also called antagonist or agonist, respectively) from suitably manipulated

cells or tissues. Standard methods for conducting screening assays are well understood in the art.

Examples of potential galectin 11 antagonists include antibodies or, in some cases, oligonucleotides or proteins which are closely related to the galectin 11 or its cognate ligand, e.g., a fragment of galectin 11 or galectin 11 ligand, or small molecules which bind to the cognate ligand, but do not elicit a response, so that the activity of the galectin 11 is prevented.

Thus in another aspect, the present invention relates to a screening kit for identifying agonists, antagonists, ligands, receptors, substrates, enzymes, etc. for galectin 11 polypeptides; or compounds which decrease or enhance the production of galectin 11, which comprises:

(a) a galectin 11 polypeptide of the invention, such as, for example, that of SEQ ID NO:2;

(b) a cell expressing a galectin 11 ligand, such as, for example, a T-cell;

(c) a cell membrane expressing a galectin 11 ligand, preferably a membrane of a T-cell;

(d) a compound containing a β galactoside sugar; or

(e) antibody to a galectin 11 polypeptide of the invention, preferably that of SEQ ID NO: 2.

It will be appreciated that in any such kit, (a), (b), (c), (d), or (e) may comprise a substantial component.

Compounds identified via assays such as those described herein, may be useful, for example, in elaborating the biological function of the galectin 11 gene product and for regulating cell growth, cell proliferation and differentiation, and apoptosis. For example, antibodies against galectin 11 and galectin 11 polypeptides, fragments, derivatives, variants or analogs of the invention may be employed to suppress galectin 11 activity to treat abnormalities resulting from elevated galectin 11. The combination of these identified compounds with a pharmaceutically acceptable carrier (e.g., as described herein) and their administration to treat or prevent growth regulatory and immunomodulatory disorders, including, but not limited to, autoimmune diseases, cancer, and inflammatory diseases, are also encompassed by the invention.

Prophylactic and Therapeutic Methods

It is to be understood that although the following discussion is specifically directed to human patients, the teachings are also applicable to any animal that expresses galectin 11.

As noted above, galectin 11 shares significant homology with other galectins. Additionally, as disclosed herein, galectin 11, like galectin 1 induces apoptosis of T-cell lines. Further, as discussed above, galectin 1 has been demonstrated to play a role in regulating cell proliferation and some immune functions (e.g., therapeutic activity against autoimmune diseases in experimental myasthenia gravis and experimental autoimmune encephalomyelitis animal model systems). Thus, it is likely that galectin 11, like galectin 1, is active in modulating growth regulatory activities (e.g., cell differentiation and/or cell proliferation) , immunomodulatory activity, cell-cell and cell-substrate interactions, and apoptosis.

Apoptosis, or programmed cell death, is a physiological mechanism involved in the deletion of peripheral T lymphocytes of the immune system, and its dysregulation can lead to a number of different pathogenic processes. Diseases associated with increased cell survival, or the inhibition of apoptosis, include cancers (such as follicular lymphomas, carcinomas with p53 mutations, and hormone-dependent tumors, such as breast cancer, prostrate cancer, Kaposi's sarcoma and ovarian cancer); autoimmune disorders (such as systemic lupus erythematosus, myasthenia gravis, and immune-related glomerulonephritis rheumatoid arthritis) and viral infections (such as herpes viruses, pox viruses and adenoviruses), inflammation; graft vs. host disease, acute graft rejection, and chronic graft rejection. Diseases associated with increased apoptosis include AIDS; neurodegenerative disorders (such as Alzheimer's disease, Parkinson's disease, Amylotrophic lateral sclerosis, Retinitis pigmentosa, Cerebellar degeneration); myelodysplastic syndromes (such as aplastic anemia), ischemic injury (such as that caused by myocardial infarction, stroke and reperfusion injury), toxin-induced liver disease (such as that caused by alcohol), septic shock, cachexia and anorexia.

Any method which neutralizes or enhances galectin 11 activity can be used to modulate growth regulatory activities (e.g., cell proliferation), immunomodulatory activity, cell-cell and cell-substrate interactions, and apoptosis.

Galectin 11 polypeptides or polynucleotides (including galectin 11 fragments, variants, derivatives, and anaologs, and galectin 11 agonists and antagonists as described herein) may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins),

or infectious. Moreover, galectin 11 polynucleotides or polypeptides can be used as a marker or detector of a particular immune system disease or disorder.

Galectin 11 polynucleotides or polypeptides (including galectin 11 fragments, variants, derivatives, and anaologs, and galectin 11 agonists and antagonists as described herein) may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. As further discussed below, galectin 11 polypeptides or polynucleotides could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocytc bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, galectin 11 polypeptides or polynucleotides (including galectin 11 fragments, variants, derivatives, and anaologs, and galectin 11 agonists and antagonists as described herein) can also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, galectin 11 polynucleotides or polypeptides could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, galectin 11 polynucleotides or polypeptides that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting, important in the treatment of heart attacks (infarction), strokes, or scarring.

Galectin 11 polynucleotides or polypeptides (including galectin 11 fragments, variants, derivatives, and anaologs, and galectin 11 agonists as described herein) may also be useful in treating or detecting autoimmune disorders. As disclosed herein, galectin 11 induces apoptosis of T-cell lines (see Example 5, Figures 5A and 5B). Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of galectin 11 polypeptides or polynucleotides that can inhibit an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by galectin 11 include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid

syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitis, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by galectin 11 polypeptides or polynucleotides. Moreover, galectin 11 can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

Galectin 11 polynucleotides or polypeptides (including galectin 11 fragments, variants, derivatives, and anaologs, and galectin 11 agonists as described herein) may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of galectin 11 polypeptides or polynucleotides that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, galectin 11 polypeptides or polynucleotides (including galectin 11 fragments, variants, derivatives, and anaologs, and galectin 11 agonists as described herein) may also be used to modulate inflammation. For example, galectin 11 polypeptides or polynucleotides may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1).

Galectin 11 polypeptides or polynucleotides (including galectin 11 fragments, variants, derivatives, and anaologs, and galectin 11 agonists and antagonists as described herein) can be used to treat or detect hyperproliferative disorders, including neoplasms. Galectin 11 polypeptides or polynucleotides may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, galectin 11

polypeptides or polynucleotides may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by galectin 11 polynucleotides or polypeptides include, but are not limited to, neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, prostate, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testiclcs, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, splecn, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by galectin 11 polynucleotides or polypeptides. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstron's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

Galectin 11 polypcptides or polynucleotides (including galectin 11 fragments, variants, derivatives, and anaologs, and galectin 11 agonists and antagonists as described herein) can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, galectin 11 polypeptides or polynucleotidcs may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by galectin 11 polynucleotides or polypeptides. Examples of viruses, include, but are not limited to, the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes

Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiollitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox , hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted discases, skin discases (c.g., Kaposi's, warts), and viremia. Galectin 11 polypeptides or polynucleotides can be used to treat or detcct any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by galectin 11 polynucleotides or polypeptides include, but are not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonclla, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Menigococcal), Pasteurellacea Infections (e.g., Actinobacillus, Heamophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. galectin 11 polypeptides or polynucleotides can be used to treat or detect any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by galectin 11 polynucleotides or polypeptides include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (c.g., dysentcry, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. Galectin 11 polypeptides or polynucleotides can be used to treat or detect any of these symptoms or discases.

Galectin 11 polynucleotides or polypeptides (including galectin 11 fragments, variants, derivatives, and anaologs, and galectin 11 agonists and antagonists as described herein) can be used to differentiate, proliferate, and attract ccIls, leading to the regeneration of tissues. (Sec, Scicncc 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissuc damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteocarthritis, periodontal disease, liver failurc), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestinc, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vascular (including vascular endothelium), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regencration also may include angiogenesis.

Moreover, galectin 11 polynucleotides or polypeptides (including galectin 11 fragments, variants, derivatives, and anaologs, and galectin 11 agonists and antagonists as described herein) may increase regeneration of tissucs difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. galectin 11 polynucleotides or polypeptides of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defccts. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using galectin 11 polynucleotides or polypeptides to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord

disorders, head trauma, cerebrovascular discase, and stoke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the galectin 11 polynucleotides or polypeptides.

Thus, in one aspect, the present invention is directed to a method for enhancing apoptosis, cell proliferation, cell differentiation, or other cell growth activity regulated by galectin 11, which involves administering to an individual in need of an increased level of galectin 11 functional or biological activity, a therapcutically effective amount of galectin 11 polypeptide, fragment, variant, derivative, or analog, or an agonist capable of increasing galectin 11 mediated cellular responses. In specific embodiments, galectin-11 mediated signaling is increased to treat a disease wherein decreased apoptosis is exhibited.

Given the activities modulated by galectin 11, it is readily apparent that a substantially altered (increased or decreased) level of expression of galectin 11 in an individual compared to the standard or "normal" level produces pathological conditions such as those described above. It will also be appreciated by one of ordinary skill that the galectin 11 polypeptides of the invention will exert its modulating activities on any of its target cells. Therefore, it will be appreciatcd that conditions caused by a decrease in the standard or normal level of galectin 11 activity in an individual, can be treated by administration of galectin 11 protein or an agonist thereof.

In addition to treating diseases associated with elevated or decreased levels of galectin 11 activity, the invention encompasses mcthods of administering galectin 11 polypeptides or polynucleotides (including fragments, variants, derivatives and analogs, and agonists and antagonists as described herein) to elevate galectin 11 associated biological activity.

For example, any method which clevates galectin 11 concentration and/or activity can be used to stimulate hematopoiesis. Using these methods, the galectin 11 polypeptide and nucleotide sequences described herein may be used to stimulate hematopoiesis. In a specific embodiment, galectin 11 polypeptides and polynucleotides are used in erythropoietin therapy, which is directed toward supplementing the oxygen carrying capacity of blood. Galectin 11 treatment within the scope of the invention includes, but is not limited, to patients generally requiring blood transfusions, such as, for example, trauma victims, surgical patients, dialysis patients, and patients with a variety of blood composition-affecting disorders, such as hemophilia, cystic fibrosis, pregnancy, menstrual disorders, early anemia of prematurity, spinal cord injury, space

flight, aging, various neoplastic disease states, and the like. Examples of patient conditions that require supplementation of the oxygen carrying capacity of blood and which are within the scope of this invention, include but are not limited to: treatment of blood disorders characterized by low or defective red blood cell production, anemia associated with chronic renal failure, stimulation of reticulocyte response, development of ferrokinetic effects (such as plasma iron turnover effects and marrow transit time effects), erythrocyte mass changes, stimulation of hemoglobin C synthesis, and increasing levels of hematocrit in vertebrates. The invention also provides for treatment to enhance the oxygen-carrying capacity of an individual, such as for example, an individual encountering hypoxic environmental conditions.

The invention also encompasses combining the galectin 11 polypeptides and polynucleotides described herein with other proposed or conventional hematopoietic therapies. Thus, for example, galectin 11 can be combined with compounds that singly exhibit erythropoietic stimulatory effects, such as erythropoietin, testosterone, progenitor cell stimulators, insulin-like growth factor, prostaglandins, serotonin, cyclic AMP, prolactin, and triiodothyzonine. Also encompassed are combinations with compounds generally used to treat aplastic anemia, such as methenolene, stanozolol, and nandrolone; to treat iron-deficiency anemia, such as iron preparations; to treat malignant anemia, such as vitamin B12 and/or folic acid; and to treat hemolytic anemia, such as adrenocortical steroids, e.g., corticoids. See e.g., Resegotti et al., 1981, Panminerva Medica, 23:243-248; Kurtz, 1982, FEBS Letters, 14a:105-108; McGonigle et al., 1984, Kidney Int., 25:437-444; and Pavlovic-Kantera, 1980, Expt. Hematol., 8(supp. 8) 283-291.

Compounds that enhance the effects of or synergize with erythropoietin are also useful as adjuvants herein, and include but are not limited to, adrenergic agonists, thyroid hormones, androgens, hepatic erythropoietic factors, erythrotropins, and erythrogenins, See for e.g., Dunn, "Current Concepts in Erythropoiesis", John Wiley and Sons (Chichester, England, 1983); Weiland et al., 1982, Blut, 44:173-175; Kalmani, 1982, Kidney Int., 22:383-391; Shahidi, 1973, New Eng. J. Med., 289:72-80; Urabe et al., 1979, J. Exp. Med., 149:1314-1325; Billat et al., 1982, Expt. Hematol., 10:133-140; Naughton et al., 1983, Acta Haemat, 69:171-179; Cognote et al. in abstract 364, Proceedings 7th Intl. Cong. of Endocrinology (Quebec City, Quebec, July 1-7, 1984); and Rothman et al., 1982, J. Surg. Oncol., 20:105-108.

Methods for stimulating hematopoiesis comprise administering a hematopoietically effective amount (i.e, an amount which effects the formation of blood cells) of a pharmaceutical composition containing galectin 11 to a patient. The galectin 11 is administered to the patient by any suitable technique, including but not limited to,

parenteral, sublingual, topical, intrapulmonary and intranasal, and those techniques further discussed herein. The pharmaceutical composition optionally contains one or more members of the group consisting of erythropoietin, testosterone, progenitor cell stimulators, insulin-like growth factor, prostaglandins, serotonin, cyclic AMP, prolactin, triiodothyzonine, methenolene, stanozolol, and nandrolone, iron preparations, vitamin B12, folic acid and/or adrenocortical steroids. The galectin 11 and cotreatment drug(s) are suitably delivered by separate or by the same administration route, and at the same or at different times, depending, e.g., on dosing, the clinical condition of the patient, etc.

For treating abnormal conditions related to an under-expression of galectin 11 and its activity, or in which elevated or decreased levels of galectin 11 are desired, several approaches are available. One approach comprises administering to an individual in need of an increased level of galectin 11 in the body, a therapeutically effective amount of an isolated galectin 11 polypeptide, fragment, variant, derivative or analog of the invention, or a compound which activates galectin 11, i.e., an agonist as described above, optionally in combination with a pharmaceutically acceptable carrier. Alternatively, gene therapy may be employed to effect the endogenous production of galectin 11 by the relevant cells in the subject. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector using techniques known in the art. The retroviral expression construct may then be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a subject for engineering cells *in vivo* and expression of the polypeptide *in vivo*. For a overview of gene therapy, see Chapter 20, Gene Therapy and other Molecular Genetic-based Therapeutic Approaches, (and references cited therein) in Human Molecular Genetics, T Strachan and A P Read, BIOS Scientific Publishers Ltd (1996).

Further, treatment can be administered, for example, in the form of gene replacement therapy. Specifically, one or more copies of a galectin 11 nucleotide sequence of the invention that directs the production of a galectin 11 gene product exhibiting normal function, may be inserted into the appropriate cells within a patient or animal subject, using vectors which include, but are not limited to, adenovirus, adeno-associated virus, retrovirus and herpesvirus vectors, in addition to other particles that introduce DNA into cells, such as liposomes and gene activated matrices. Because the galectin 11 gene is expressed in neutrophils, such gene replacement techniques should be capable of delivering galectin 11 gene sequence to these cells within patients, or,

alternatively, should involve direct administration of such galectin 11 polynucleotide sequences to the site of the cells in which the galectin 11 gene sequences are to be expressed. Alternatively, targeted homologous recombination can be utilized to correct the defective endogenous galectin 11 gene and/or regulatory sequences thereof (e.g., promoter and enhancer sequences), or alternatively, to "turn on" other dormant galectin 11 activity in the appropriate tissue or cell type.

Additional methods which may be utilized to increase the overall level of galectin 11 expression and/or galectin 11 activity include the introduction of appropriate galectin 11-expressing cells, preferably autologous cells, into a patient at positions and in numbers which are sufficient to ameliorate the symptoms of abnormalities in cells growth regulation. Such cells may be either recombinant or non-recombinant. Among the cells which can be administered to increase the overall level of galectin 11 gene expression in a patient are normal cells, which express the galectin 11 gene. Cell-based gene therapy techniques are well known to those skilled in the art, see, e.g., Anderson et al., U.S. Patent No. 5,399,349; and Mulligan & Wilson, U.S. Patent No. 5,460,959.

If the activity of galectin 11 is in excess, several approaches are available to reduce or inhibit galectin 11 activity using molecules derived from the polypeptide and polynucleotide sequences described above. Accordingly, a further aspect of the invention is related to a method for treating an individual in need of a decreased level of galectin 11 activity in the body comprising, administering to such an individual a composition comprising a therapeutically effective amount of a galectin 11 polypeptide, fragment, variant, derivative or analog of the invention which acts as a galectin 11 antagonist, optionally, in combination with a pharmaceutically acceptable carrier. Preferably, galectin 11 activity is decreased to treat a disease wherein increased apoptosis or other cell growth activity regulated by galectin 11 is exhibited. Polypeptides, derivatives, variants and analogs of the invention which function as antagonists of galectin 11 can routinely be identified using the assays described *infra* and other techniques known in the art. Preferred antagonists for use in the present invention are galectin 11-specific antibodies.

Thus, one embodiment of the invention comprises administering to a subject an inhibitor compound (antagonist), such as for example, an antibody or fragment, variant, derivative or analog of the invention, along with a pharmaceutically acceptable carrier in an amount effective to suppress (i.e. lower) galectin 11 activity.

In another approach, galectin 11 activity can be reduced or inhibited by decreasing the level of galectin 11 gene expression. In one embodiment, this is accomplished through the use of antisense sequences, either internally generated or

separately administered (see, for example, O'Connor, J. Neurochem. (1991) 56:560 in Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988). Antisense technology can be used to control gene expression through antisense DNA or RNA or through triple-hclix formation. Antisense techniques are discussed, for example, in Okano, J. Neurochem. 56:560 (1991); *Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression*, CRC Press, Boca Raton, FL (1988). Triple helix formation is discussed in, for instance, Lee et al., Nucleic Acids Research 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991). The methods are based on binding of a polynucleotide to a complementary DNA or RNA. For example, the 5' coding portion of a polynucleotide that encodes galectin 11 polypeptidc of the present invention may be used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription thereby preventing transcription and thc production of the galectin 11 polypeptide. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of thc mRNA molecule into polypeptide.

In one embodiment, the galectin 11 antisense nucleic acid of the invention is produccd intracellularly by transcription from an exogenous sequence. For example, a vector or a portion thereof, is transcribed, producing an antisense nucleic acid (RNA) of the invention. Such a vector would contain a sequence encoding the galectin 11 antisense nucleic acid. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others know in the art, used for replication and expression in vertebrate cells. Expression of the sequence encoding galectin 11, or fragments thereof, can be by any promoter known in the art to act in vertebrate, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include, but are not limited to, the SV40 early promoter region (Bernoist and Chambon, Nature 29:304-310 (1981), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., Cell 22:787-797 (1980), the herpes thymidine promoter (Wagner et al., Proc. Natl. Acad. Sci. U.S.A. 78:1441-1445 (1981)), the regulatory scquences of the metallothionein gene (Brinster et al., Nature 296:39-42 (1982)), etc.

The antisense nucleic acids of the invention comprise a sequence complementary to at least a portion of an RNA transcript of a galectin 11 gene. However, absolute complementarity, although preferred, is not requircd. A sequence "complementary to at least a portion of an RNA," referred to herein, means a sequence having sufficient

complementarity to be able to hybridize with the RNA, forming a stable duplex; in the case of double stranded galectin 11 antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the larger the hybridizing nucleic acid, the more base mismatches with a galectin 11 RNA it may contain and still form a stable duplex (or triplex as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

Potential galectin 11 antagonists according to the invention also include catalytic RNA, or a ribozyme (See, e.g., PCT International Publication WO 90/11364, published October 4, 1990; Sarver et al., Science 247:1222-1225 (1990). While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy galectin 11 mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Haseloff and Gerlach, Nature 334:585-591 (1988). There are numerous potential hammerhead ribozyme cleavage sites within the nucleotide sequence of galectin 11 (FIG.1; SEQ ID NO:1). Preferably, the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the galectin 11 mRNA; i.e., to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts. DNA constructs encoding the ribozyme may be introduced into the cell in the same manner as described above for the introduction of antisense encoding DNA. Since ribozymes, unlike antisense molecules are catalytic, a lower intracellular concentration is required for efficiency.

Endogenous galectin 11 gene expression can also be reduced by inactivating or "knocking out" the galectin 11 gene or its promoter using targeted homologous recombination (e.g., see Smithies et al., Nature 317:330-234 (1985); Thomas et al., Cell 51:503-512 (1987); Thompson et al., Cell 5:313-321 (1989); each of which is incorporated by reference herein in its entirety). Such approach can be adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors.

Alternatively, endogenous galectin 11 gene expression can be reduced by targeted deoxyribonucleotide sequences complementary to the regulatory region of the galectin 11 gene (i.e., the galectin 11 promoter and/or enhancers) to form triple helical

structures that prevent transcription of the galectin 11 gene in target cells in the body, see generally, Helene et al., Ann, N.Y. Acad. Sci. 660:27-36 (1992); Helene, C., Anticancer Drug Des., 6(6):569-584 (1991); and Maher, L.J., Bioassays 14(12):807-815 (1992)).

In yet another embodiment of the invention, the activity of galectin 11 can be reduced using a "dominant negative". To this end, constructs which encode defective galectin 11, such as, for example, mutants lacking all or a portion of region of galectin 11 that binds β-galactosides, can be used in gene therapy approaches to diminish the activity of galectin 11 on appropriate target cells. For example, nucleotide sequences that direct host cell expression of galectin 11 in which all or a portion of the region of galectin 11 that binds β-galactoside is altered or missing can be introduced into neutrophil cells, or other cells or tissue which express galectin 11 (either by *in vivo* or *ex vivo* gene therapy methods as for example, described herein). Alternatively, targeted homologous recombination can be utilized to introduce such deletions or mutations into the subjects endogenous galectin 11 gene in neutrophils or other cells expressing galectin 11.

Formulation and administration

It will be appreciated that conditions caused by a decrease in the standard or normal level of galectin 11 activity in an individual, can be treated by administration of galectin 11 polypeptide or fragment, variant, derivative, or analog of the invention or an agonist thereof. Thus, the invention further provides a method of treating an individual in need of an increased level of galectin 11 activity comprising administering to such an individual a pharmaceutical composition comprising an effective amount of an isolated galectin 11 polypeptide or fragment, variant, derivative, or analog of the invention, such as for example, the full length form of the galectin 11, effective to increase the galectin 11 activity level in such an individual.

The dosage range required depends on the choice of peptide, the route of administration, the nature of the formulation, the nature of the subject's condition, and the judgment of the attending practitioner. As a general proposition, the total pharmaceutically effective amount of galectin 11 polypeptide administered parenterally per dose will be in the range of about 1 μg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans this dose is in the range of 0.1-100 mg/kg of subject, or between about 0.01 and 1 mg/kg/day. If given continuously, the galectin 11 polypeptide is typically administered

at a dose rate of about 1 µg/kg/hour to about 50 µg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. Wide variations in the needed dosage, however, are to be expected in view of the variety of compounds available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted using standard empirical routines for optimization, as is well understood in the art.

Pharmaceutical compositions containing the galectin 11 polypeptidcs and polynucleotides of the invention (including fragments, variants, derivatives or analogs), and galectin 11 agonists and antagonists may be routinely formulated in combination with a pharmaceutically acceptable carrier. By "pharmaceutically acceptable carrier" is meant a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. In a specific embodiment, "pharmaceutically acceptable" means approved by a regulatory agency of the federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly humans. Nonlimiting examples of suitable pharmaceutical carriers according to this embodiment are provided in "Remington's Pharmaceutical Sciences" by E.W. Martin, and include sterile liquids, such as water, saline, buffered saline, glycerol, ethanol, and oils, including thosc of pctroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Formulation should suit the mode of administration, and is well within the skill of the art. For example, water is a preferred carrier when the pharmaceutical composition is administered intravenously. Saline solutions and aqueous dextrose and glycerol solutions can be employed as liquid carriers, particularly for injectable solutions. The invention additionally rclates to pharmaceutical packs and kits comprising one or morc containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

Polypeptides and other compounds of the present invention may be administered alone or in conjunction with other compounds, such as therapeutic compounds. The pharmaceutical composition of the invention may be administered orally, rectally, parenterally, intracistemally, intravaginally, intraperitoneally, topically (as by powders, ointments, drops or transdermal patch), bucally, or as an oral or nasal spray. Preferred forms of systemic administration of the pharmaceutical compositions include parenteral injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, intrasternal, intraarticular or intraperitoneal, can be uscd. Altcmative mcans for systemic administration include transmucosal and

transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if properly formulated in enteric or encapsulated formulations, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels and the like.

Polypeptides used in treatment can also be generated endogenously in the subject, in treatment modalities often referred to as "gene therapy" as described above. Thus, for example, cells from a subject may be engineered with a polynucleotide, such as a DNA or RNA, to encode a polypeptide ex vivo, and for example, by the use of a retroviral plasmid vector. The cells are then introduced into the subject.

Chromosome Assays

The nucleic acid molecules of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on human chromosome 11. The mapping of DNAs to chromosomes according to the present invention is an important first step in correlating those sequences with genes associated with disease.

Since the galectin 11 gene has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, *Mendelian Inheritance In Man*, available on-line through Johns Hopkins University, Welch Medical Library. The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

Next, it is necessary to determine the differences in the cDNA or genomic sequence between affected and unaffected individuals. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

Examples

Example 1: Expression and Purification of Galectin 11 in E. coli

The DNA sequence encoding the galectin 11 protein in the deposited cDNA clone is amplified using PCR oligonucleotide primers specific to the amino terminal sequences of the galectin 11 protein and to vector sequences 3' to the gene. Additional nucleotides containing restriction sites to facilitate cloning are added to the 5' and 3' sequences respectively.

The 5' galectin 11 oligonucleotide primer has the sequence 5' cgc CCATGG ATGAGCCCCAGGCTGGAGGTG 3' (SEQ ID NO:5) containing the underlined NcoI restriction site and nucleotides 49 to 69 of the galectin 11 nucleotide sequence depicted in Figure 1 (SEQ ID NO:1).

The 3' galectin 11 primer has the sequence 5' cgc AAGCTT TCAGGAGTGGACACAGTAG 3' (SEQ ID NO:6) containing the underlined HindIII restriction site followed by nucleotides complementary to position 431 to 451 of the galectin 11 nucleotide sequence depicted in Figure 1 (SEQ ID NO:1).

The restriction sites are convenient to restriction enzyme sites in the bacterial expression vector pQE60 which are used for bacterial expression in these examples. (Qiagen, Inc. 9259 Eton Avenue, Chatsworth, CA, 91311). pQE60 encodes ampicillin antibiotic resistance ("Amp^r") and contains a bacterial origin of replication ("ori"), an IPTG inducible promoter, a ribosome binding site ("RBS"), a 6-His tag and restriction enzyme sites.

The amplified galectin 11 DNA and the pQE60 vector is digested with NcoI and HindIII and the digested DNAs are then ligated together. Insertion of the galectin 11 polypeptide DNA into the restricted pQE60 vector places the galectin 11 polypeptide coding region downstream of and operably linked to the vector's IPTG-inducible promoter and in-frame with an initiating AUG appropriately positioned for translation of galectin 11.

The ligation mixture is transformed into competent *E. coli* cells using standard procedures. Such procedures are described in Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989). *E. coli* strain M15/rep4, containing multiple copies of the plasmid pREP4, which expresses lac repressor and confers kanamycin resistance ("Kan^r"), is used in carrying out the example described herein. This strain, which is only one of many that are suitable for expressing galectin 11 protein, is available commercially from Qiagen.

Transformants are identified by their ability to grow on LB plates in the presence of ampicillin and kanamycin. Plasmid DNA is isolated from resistant colonies and the identity of the cloned DNA confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both ampicillin (100 µg/ml) and kanamycin (25 µg/ml).

The O/N culture is used to inoculate a large culture, at a dilution of approximately 1:100 to 1:250. The cells are grown to an optical density at 600nm ("OD600") of between 0.4 and 0.6. Isopropyl-B-D-thiogalactopyranoside (IPTG) is then added to a final concentration of 1 mM to induce transcription from *lac* repressor sensitive promoters, by inactivating the *lac*I repressor. Cells subsequently are incubated further for 3 to 4 hours. Cells then are harvested by centrifugation and disrupted, by standard methods. Inclusion bodies are purified from the disrupted cells using routine collection techniques, and protein is solubilized from the inclusion bodies into 8M urea. The 8M urea solution containing the solubilized polypeptide is passed over a PD-10 column in 2X phosphate-buffered saline ("PBS"), thereby removing the urea, exchanging the buffer and refolding the protein. The polypeptide is purified by a further step of chromatography to remove endotoxin. Then, it is sterile filtered. The sterile filtered protein preparation was stored in 2X PBS at a concentration of 95 µ/ml.

Example 2: Cloning and Expression of Galectin 11 protein in a Baculovirus Expression System

The cDNA sequence encoding the full length galectin 11 protein in the deposited clone is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene:

The 5' galectin 11 oligonucleotide primer has the sequence 5' cgc CCC GGG GCCT ATGAGCCCCAGGCTGGAGG 3' (SEQ ID NO:7) containing the underlined SmaI restriction site and nucleotides 49 to 66 of the galectin 11 nucleotide sequence depicted in Figure 1 (SEQ ID NO:1).

The 3' Galectin 11 primer has the sequence 5' cgc GGT ACC TCAGGAGTGGACACAGTAG 3' (SEQ ID NO:8) containing the underlincd Asp718 restriction site followed by nucleotides complementary to position 432 to 450 of the galectin 11 nucleotide sequence depicted in Figure 1 (SEQ ID NO:1).

An efficient signal for initiation of translation in eukaryotic cells, as described by Kozak, M., J. Mol. Biol. 196: 947-950 (1987) is appropriately located in the vector portion of the construct.

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with XbaI and again is purified on a 1% agarose gel. This fragment is designated herein F2.

The vector pA2-GP is used to express the galectin 11 protein in the baculovirus expression system, using standard methods, as described in Summers et al, A MANUAL OF METHODS FOR BACULOVIRUS VECTORS AND INSECT CELL CULTURE PROCEDURES, Texas Agricultural Experimental Station Bulletin No. 1555 (1987). This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites. The signal peptide of AcMNPV gp67, including the N-terminal methionine, is located just upstream of a BamHI site. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For an easy selection of recombinant virus the beta-galactosidase gene from *E. coli* is inserted in the same orientation as the polyhedrin promoter and is followed by the polyadenylation signal of the polyhedrin gene. The polyhedrin sequences are flanked at both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate viable virus that express the cloned polynucleotide.

Many other baculovirus vectors could be used in place of pA2-GP, such as pAc373, pVL941 and pAcIM1 provided, as those of skill readily will appreciate, that construction provides appropriately located signals for transcription, translation, trafficking and the like, such as an in-frame AUG and a signal peptide, as required. Such vectors are described in Luckow et al., Virology 170: 31-39, among others.

The plasmid is digested with the restriction enzyme SmaI and Asp718 and then is dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.). This vector DNA is designated herein V2 .

Fragment F2 and the dephosphorylated plasmid V2 are ligated together with T4 DNA ligase. *E. coli* HB101 cells are transformed with ligation mix and spread on culture plates. Bacteria are identified that contain the plasmid with the human galectin 11 gene by digesting DNA from individual colonies using XbaI and then analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing. This plasmid is designated herein pBacgalectin 11.

5 μg of the plasmid pBacgalectin 11 is co-transfected with 1.0 μg of a commercially available linearized baculovirus DNA ("BaculoGold baculovirus DNA", Pharmingen, San Diego, CA.), using the lipofection method described by Felgner et

al., Proc. Natl. Acad. Sci. USA 84: 7413-7417 (1987). 1µg of BaculoGold virus DNA and 5 µg of the plasmid pBacgalectin 11 are mixed in a sterile well of a microtiter plate containing 50 µl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards 10 µl Lipofectin plus 90 µl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is rocked back and forth to mix the newly added solution. The plate is then incubated for 5 hours at 27°C. After 5 hours the transfection solution is removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. The plate is put back into an incubator and cultivation is continued at 27°C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, cited above. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10).

Four days after serial dilution, the virus is added to the cells. After appropriate incubation, blue stained plaques are picked with the tip of an Eppendorf pipette. The agar containing the recombinant viruses is then resuspended in an Eppendorf tube containing 200 µl of Grace's medium. The agar is removed by a brief centrifugation and the supernatant containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4°C. A clone containing properly inserted hESSB I, II and III is identified by DNA analysis including restriction mapping and sequencing. This is designated herein as V-galectin 11.

Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus V-galectin 11 at a multiplicity of infection ("MOI") of about 2 (about 1 to about 3). Six hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Gaithersburg). 42 hours later, 5 µCi of ^{35}S-methionine and 5 µCi ^{35}S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then they are harvested by centrifugation, lysed and the labeled proteins are visualized by SDS-PAGE and autoradiography.

Example 3: Cloning and Expression in Mammalian Cells

Most of the vectors used for the transient expression of the galectin 11 polypeptide genc sequence in mammalian cells should carry the SV40 origin of replication. This allows the replication of the vector to high copy numbers in cells (e.g. COS cells) which express thc T antigen required for the initiation of viral DNA synthesis. Any other mammalian cell line can also be utilized for this purpose.

A typical mammalian expression vector contains the promoter element, which mediates the initiation of transcription of mRNA, the protein coding sequence, and signals required for the termination of transcription and polyadenylation of thc transcript. Additional clcmcnts include cnhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription can be achieved with the early and late promoters from SV40, thc long tcrminal repcats (LTRs) from Retroviruses, e.g. RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular signals can also be used (e.g. human actin promoter). Suitablc cxpression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146) and pBC12MI (ATCC 67109). Mammalian host cells that could be used include, human Hela, 283, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, African green monkey cells, quail QC1-3 cells, mouse L cells and Chinese hamster ovary cells.

Altcmatively, thc gene can bc expressed in stable cell lines that contain the gene integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) is a useful marker to devclop cell lines that carry several hundred or even several thousand copies of the gene of interest. Another uscful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992)). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) cells are often used for the production of proteins.

The expression vectors pC1 and pC4 contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-4470 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., Cell 41:521-530 (1985)). Multiple cloning sites, e.g. with the restriction enzyme cleavage sites BamHI, XbaI and

Asp718, facilitate the cloning of the gene of interest. The vectors contain in addition the 3 intron, the polyadenylation and termination signal of the rat preproinsulin gene.

Example 3(a): Cloning and Expression in COS Cells

The expression plasmid, pgalectin 11, is made by cloning a cDNA encoding galectin 11 into the expression vector pcDNAI/Amp (which can be obtained from Invitrogen, Inc.).

The expression vector pcDNAI/amp contains: (1) an *E. coli* origin of replication effective for propagation in *E. coli* and other prokaryotic cells; (2) an ampicillin resistance gene for selection of plasmid-containing prokaryotic cells; (3) an SV40 origin of replication for propagation in eukaryotic cells; (4) a CMV promoter, a polylinker, an SV40 intron, and a polyadenylation signal arranged so that a cDNA conveniently can be placed under expression control of the CMV promoter and operably linked to the SV40 intron and the polyadenylation signal by means of restriction sites in the polylinker.

A DNA fragment encoding the galectin 11 protein and an HA tag fused in frame to its 3' end is cloned into the polylinker region of the vector so that recombinant protein expression is directed by the CMV promoter. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein described by Wilson *et* al., Cell 37: 767 (1984). The fusion of the HA tag to the target protein allows easy detection of the recombinant protein with an antibody that recognizes the HA epitope.

The plasmid construction strategy is as follows. The galectin 11 cDNA of the deposited clone is amplified using primers that contain convenient restriction sites, much as described above regarding the construction of expression vectors for expression of galectin 11 in *E. coli*. To facilitate detection, purification and characterization of the expressed galectin 11, one of the primers contains a hemagglutinin tag ("HA tag") as described above.

Suitable primers include the following, which are used in this example. The 5' galectin 11 primer has the sequence 5' cgc CCC GGG gcc atc ATG GCCTATC ATGAGCCCCAGGCTGGAGG 3' (SEQ ID NO:9) containing the underlined SmaI restriction enzyme site followed by nucleotide sequence 49 to 66 of Figure 1 (SEQ ID NO:1).

The 3' galectin 11 primer has the sequence 5' cgc GGT ACC TCAGGAGTGGACACAGTAG 3' (SEQ ID NO:8) containing the Asp718 restriction followed by nucleotides complementary to nucleotides 432 to 450 of the galectin 11 nucleotide sequence depicted in Figure 1 (SEQ ID NO:1).

The PCR amplified DNA fragment and the vector, pcDNAI/Amp, are digested with HindIII and XhoI and then ligated. The ligation mixture is transformed into *E.*

coli strain SURE (available from Stratagenc Cloning Systems, 11099 North Torrey Pines Road, La Jolla, CA 92037), and the transformed culture is plated on ampicillin media plates which then are incubated to allow growth of ampicillin resistant colonies. Plasmid DNA is isolated from resistant colonies and examined by restriction analysis and gel sizing for the presence of the galectin 11-encoding fragmcnt.

For expression of recombinant galectin 11, COS cells are transfected with an expression vector, as described above, using DEAE-DEXTRAN, as described, for instance, in Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, Cold Spring Laboratory Press, Cold Spring Harbor, New York (1989). Cells are incubated under conditions for expression of galectin 11 by the vector.

Expression of the galectin 11 HA fusion protein is detected by radiolabelling and immunoprecipitation, using methods described in, for example Harlow et al., ANTIBODIES: A LABORATORY MANUAL, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1988). To this end, two days after transfection, the cells are labeled by incubation in media containing ^{35}S-cysteine for 8 hours. The cells and the media are collected, and the cells are washed and the lysed with detergent-containing RIPA buffer: 150 mM NaCl, 1% NP-40, 0.1% SDS, 1% NP-40, 0.5% DOC, 50 mM TRIS, pH 7.5, as described by Wilson et al. cited above. Proteins are precipitated from the cell lysate and from the culture media using an HA-specific monoclonal antibody. The precipitated proteins then are analyzed by SDS-PAGE gels and autoradiography. An cxpression product of the expected size is seen in the cell lysate, which is not seen in negative controls.

Example 3(b): Cloning and Expression in CHO Cells

The vector pC1 is used for the exprcssion of galectin 11 protein. Plasmid pC1 is a derivative of the plasmid pSV2-dhfr [ATCC Accession No. 37146]. Both plasmids contain the mouse DHFR gene under control of the SV40 early promoter. Chinese hamster ovary- or other cells lacking dihydrofolate activity that are transfected with these plasmids can be selectcd by growing the cells in a selective medium (alpha minus MEM, Life Technologies) supplemented with the chemotherapeutic agent methotrexate. The amplification of the DHFR genes in cells resistant to methotrexate (MTX) has been well documented (see, e.g., Alt, F.W., Kellems, R.M., Bertino, J.R., and Schimke, R.T., 1978, J. Biol. Chem. 253:1357-1370, Hamlin, J.L. and Ma, C. 1990, Biochem. et Biophys. Acta, 1097:107-143, Page, M.J. and Sydcnham, M.A. 1991, Biotechnology Vol. 9:64-68). Cells grown in increasing concentrations of MTX develop resistance to the drug by overproducing the target enzyme, DHFR, as a result of amplification of the DHFR gene. If a second gene is linked to the DHFR gene it is

usually co-amplified and over-expressed. It is state of the art to develop cell lines carrying more than 1,000 copies of the genes. Subsequently, when the methotrexate is withdrawn, cell lines contain the amplified genc intcgrated into the chromosome(s).

Plasmid pC1 contains for the expression of the gene of interest a strong promoter of the long terminal repeat (LTR) of the Rouse Sarcoma Virus (Cullen, *et al.*, Molecular and Cellular Biology, March 1985:438-4470) plus a fragment isolated from the enhancer of the immediate early gene of human cytomegalovirus (CMV) (Boshart et al., Cell 41:521-530, 1985). Downstream of the promoter are the following single restriction enzyme cleavage sites that allow the integration of the genes: BamHI, Pvull, and Nrul. Behind these cloning sites the plasmid contains translational stop codons in all three reading frames followed by the 3 intron and the polyadenylation site of the rat preproinsulin gene. Other high efficient promoters can also be used for the expression, e.g., the human -actin promoter, the SV40 early or late promoters or the long terminal repeats from other retroviruses, e.g., HIV and HTLVI. For the polyadenylation of the mRNA other signals, e.g., from the human growth hormone or globin genes can be used as well.

Stable cell lines carrying a gene of interest integrated into the chromosomes can also be selected upon co-transfection with a selectable marker such as gpt, G418 or hygromycin. It is advantageous to use more than one selectable marker in the beginning, e.g., G418 plus methotrexate.

The plasmid pC1 is digested with the restriction enzyme BamHI and then dephosphorylated using calf intestinal phosphatase by procedures known in the art. The vector is then isolated from a 1% agarose gel.

The DNA sequence encoding galectin 11, ATCC Deposit No. 209053 is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene:

The 5' Galectin 11 primer has the sequence 5' cgc CCC GGG gcc atc ATG GCCTATC ATGAGCCCCAGGCTGGAGG 3' (SEQ ID NO:9) containing the underlined SmaI restriction enzyme site followed by nucleotide sequence 49-66 of Figure 1 (SEQ ID NO:1). Inserted into an expression vector, as described below, the 5' end of the amplified fragment encoding human galectin 11 provides an efficient signal peptide. An efficient signal for initiation of translation in eukaryotic cells, as described by Kozak, M., J. Mol. Biol. 196:947-950 (1987) is appropriately located in the vector portion of the construct.

The 3' Galectin 11 primer has the sequence 5' cgc GGT ACC TCAGGAGTGGACACAGTAG 3' (SEQ ID NO:8) containing the Asp718 restriction

followed by nucleotides complementary to nucleotides 432-450 of the galectin 11 nucleotide sequence depicted in Figure 1 (SEQ ID NO:1).

The amplified fragments are isolated from a 1% agarose gel as described above and then digested with the endonucleases SmaI and Asp718 and then purified again on a 1% agarose gel.

The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 cells are then transformed and bacteria identified that contained the plasmid pC1 inserted in the correct orientation using the restriction enzyme SmaI. The sequence of the inserted gene is confirmed by DNA sequencing.

Transfection of CHO-DHFR-cells

Chinese hamster ovary cells lacking an active DHFR enzyme are used for transfection. 5 µg of the expression plasmid C1 are cotransfected with 0.5 µg of the plasmid pSVneo using the lipofecting method (Felgner et al., supra). The plasmid pSV2-neo contains a dominant selectable marker, the gene neo from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) and cultivated from 10-14 days. After this period, single clones are trypsinized and then seeded in 6-well petri dishes using different concentrations of methotrexate (25 nM, 50 nM, 100 nM, 200 nM, 400 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (500 nM, 1 µM, 2 µM, 5 µM). The same procedure is repeated until clones grow at a concentration of 100 µM.

The expression of the desired gene product is analyzed by Western blot analysis and SDS-PAGE.

Example 4: Tissue distribution of protein expression

Northern blot analysis is carried out to examine galectin 11 gene expression in human tissues, using methods described by, among others, Sambrook *et al.*, cited above. A cDNA probe containing the entire nucleotide sequence encoding galectin 11 protein (SEQ ID NO:1) is labeled with ^{32}P using the *redi*prime DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using a CHROMA SPIN-100 column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for galectin 11 mRNA.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) are obtained from Clontech and are examined with labeled probe using ExpressHyb hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and films developed according to standard procedures.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples.

Example 5: Galectin 11 induced apoptosis in transfected cells

This example presents data demonstrating that transfection of a constitutive galectin 11 expression construct into human Jurkat T-cells induces apoptosis of the transfected cells.

A T cell is a type of lymphocyte, or "white blood cell", that mediates the cellular immune response to foreign macromolecule, termed antigens. While T cells are necessary for normal mammalian immune responses, in some instances it is desirable to inhibit their activation: for example, in some autoimmune diseases, the T cells of a subject respond to "self-antigens", i.e., macromolecule produced by the subject, rather than foreign-made macromolecule, and damage the cells and tissues of the subject. Autoimmune T cell responses are found in subjects having systemic lupus erythematosus (SLE), rheumatoid arthritis (RA), insulin-dependent diabetes, myasthenia gravis, and multiple sclerosis (MS) and contribute to the pathophysiology of each. T cells also cause graft rejection and graft versus host disease (GVHD). Graft rejection is caused by an immune response against the transplanted tissues (the graft), which are recognized as "foreign" by T cells of the recipient (host). Graft versus host disease is caused by engrafted T cells, which recognize host-made macromolecule as "foreign."

Methods

The DNA sequence encoding the galectin 11 protein in the deposited cDNA clone was amplified using PCR oligonucleotide primers specific to the amino terminal sequences of the galectin 11 protein and to vector sequences 3' to the gene. The 5' galectin 11 oligonucleotide primer had the sequence 5' CGCCGCCACCATGAGCCCCAGGC 3' (SEQ ID NO:10) containing nucleotides 49 to 61 of the galectin 11 nucleotide sequence in Figure 1 (SEQ ID NO:1). The 3' galectin 11 primer has the sequence 5' GGAATCTAGATCAGGAGTGGAC 3' (SEQ ID NO:11) containing the underlined XbaI restriction site followed by nucleotide

sequence complementary to position 439 to 450 of the galectin 11 nucleotide sequence in Figure 1 (SEQ ID NO:1).

The amplified galectin 11 fragments were isolated from a 1% agarose gel as described above, digested with the endonuclease XbaI, purified again on a 1% agarose gel, and ligated into the multiple cloning site of restricted pEF1 using T4 DNA ligase.

The pEF1 vector was generated by replacing the CMV promoter on pIRES1neo (Clontech) with the human elongation factor 1 a constitutive promoter from pEF-BOS. The EF1a promoter has been shown to be highly active in a variety of cell types (data not shown). This vector also contains a bovine growth hormone poly A signal and a ampicilin resistance gene, a ribosome binding site ("RBS"), a 6-His tag and restriction enzyme sites. This vector was digested with EcoRI, BamHI, and phosphatase using techniques known in the art.

Insertion of the isolated galectin 11 fragment into the restricted pEF1 vector placed the galectin 11 polypeptide coding region downstream of and operably associated with the vector's constitutive elongation factor-1 promoter and in-frame with an initiating AUG appropriately positioned for translation of galectin 11. *E. coli* cells were then transformed with the ligation reaction and those cells containing the desired construct (pEFLeg11) were identified using techniques known in the art. Cells containing the pEFLeg11 expression construct were then cultured under known conditions favoring high yield and the expression construct was isolated from the bacterial cell culture using techniques known in the art.

For detection of apoptosis, techniques known in the art were used to cotransfect human Jurkat T-cells with the pEFLeg11 expression construct together with a marker plasmid encoding green fluorescent protein (GFP). The transfected cells were then stained with MitoTracker Red (Molecular Probes) to determine the transmembrane potentials of mitochondria, and analyzed by two-color flow cytometry. Transfected populations were identified by emission of green fluorescence due to the expression of GFP. Apoptotic cells exhibit disrupted mitochondrial transmembrane potential and thus have lower red fluorescence emission because of their reduced ability to sequester the dye MitoTracker Red.

Results

Jurkat cells transfected with the constitutive expression plasmid for "galectin 11" underwent significant apoptosis 24h after transfection. Approximately 30% of "galectin-11" transfected cells showed reduced mitochondrial transmembrane potential compared to less than 10% in cells transfected with the control vector with no insert (pEF1) (Figure 5A).

We also followed the number of GFP positive cells during a 4-day culture period after co-transfection with either the control vector pEF1 or the "galectin-11" expression vector pEF1-Leg11. There were about 4 times more surviving GFP positive cells after transfection with pEF1 than with pEF1-Leg11 (Figure 5B).

The present invention is not to be limited in scope by the specific embodiments described herein, which are intended as single illustrations of individual aspects of the invention. Indeed, various modifications of the invention, in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and accompanying drawings. Such modifications are intended to fall within the scope of the appended claims.

The entire disclosure of all publications (including patents, patent applications, journal articlcs, laboratory manuals, books, or other documents) cited herein are hereby incorporated by reference.

Further, the Sequence Listing submitted herewith in paper and computer readable form is herein incorporated by reference in its entirely.